

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Inc.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 08:14:28 ; Search time 347 Seconds
(without alignments)
4500.606 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186

Sequence: 1 caggtgattatcctaataa.....caccacaacaccttaag 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4161431 seqs, 245089505 residues

Total number of hits satisfying chosen parameters: 8322862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.New.*
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
9: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq3:*

*pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.2	3.1	1290	6	US-10-750-185-43433
2	59.4	1.9	10968	7	US-11-075-185-35
3	59.4	1.9	78869	7	US-11-075-185-35
4	55.2	1.7	1290	6	US-10-858-730-47
5	54.8	1.7	116856	7	US-11-143-980-1
6	54.2	1.7	5679	7	US-11-075-185-36
7	54.2	1.7	14172	7	US-11-075-185-2
8	53.2	1.7	5679	7	US-11-075-185-36
9	53.2	1.7	78869	7	US-11-075-185-1
10	52.2	1.6	2302	7	US-11-110-082-20
11	52	1.6	2079	7	US-11-045-802-25
12	52	1.6	7006	6	US-10-821-234-218
13	52	1.6	18138	6	US-10-995-561-13385
14	51.6	1.6	996	6	US-10-432-483-22
15	51.6	1.6	6941	6	US-10-432-483-29
16	51.4	1.6	2148	6	US-10-821-234-29
17	51.2	1.6	3513	6	US-10-858-730-142
18	50.8	1.6	3381	7	US-11-075-185-33
19	50.8	1.6	170995	7	US-11-121-086-35
20	50.6	1.6	191684	7	US-11-121-086-2
21	50.4	1.6	171936	6	US-10-933-025-24
22	50	1.6	2033	6	US-10-995-561-55
23	50	1.6	3305	7	US-11-017-550-68

24	50	1.6	12391	7	US-11-017-550-67	Sequence 67, Appl
25	50	1.6	12591	6	US-10-995-561-13415	Sequence 13415, A
26	50	1.6	14023	6	US-10-995-561-13221	Sequence 13221, A
27	50	1.6	199130	6	US-10-995-561-13233	Sequence 13233, A
28	49.8	1.6	2584	6	US-10-821-234-836	Sequence 836, App
29	49.6	1.6	1593	6	US-10-858-730-138	Sequence 138, App
30	49.4	1.6	3408	6	US-10-858-730-40	Sequence 40, Appl
31	49.4	1.6	11070	7	US-11-075-185-34	Sequence 34, Appl
32	49.4	1.6	96128	6	US-10-995-561-13197	Sequence 13197, A
33	48.8	1.5	61718	7	US-10-995-561-13326	Sequence 13326, A
34	48.6	1.5	11119	7	US-11-143-980-11	Sequence 11, Appl
35	48.6	1.5	116856	7	US-11-143-980-1	Sequence 1, Appl1
36	48.4	1.5	4509	7	US-11-087-100-5	Sequence 5, Appl1
37	48.4	1.5	4509	7	US-11-087-084-5	Sequence 5, Appl1
38	48.4	1.5	4509	7	US-11-087-085-5	Sequence 5, Appl1
39	48.4	1.5	4982	6	US-10-276-233A-17	Sequence 17, Appl
40	48.4	1.5	11070	7	US-11-075-185-34	Sequence 34, Appl
41	48.2	1.5	1719	6	US-10-848-724-4	Sequence 4, Appl1
42	48	1.5	10968	7	US-11-075-185-35	Sequence 35, Appl
43	47.8	1.5	888	6	US-10-858-730-183	Sequence 183, App
44	47.8	1.5	944	6	US-10-667-295-150	Sequence 150, App
45	47.8	1.5	8651	6	US-10-432-483-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-10-750-185-43433
; Sequence 43433, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR IMPROVING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIORITY FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43433
; LENGTH: 1290
; TYPE: DNA
; ORGANISM: Bovine 19866881067936
US-10-750-185-43433

Query Match      3.1% Score 97.2; DB 6; Length 1290;
Best Local Similarity 75.9%; Pred. No. 6.1e-12;
Matches 120; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 993 AGCGCATGTTCTCTTTTAAAGTTTAAAGTTTCTGCTGATCCACGGCTATTAC 1052
    |||||
DB 478 AGACGATGTTCTCTTTTAAAGTTTAAAGTTTCTGCTGATCCACGGCTATTAC 537
    |||||
QY 1053 AATATTTTGTGATGATGATTTTGGCGATCCATCATCTGAGGTTTCAAGAGGCAAA 1112
    |||||
DB 538 AACGTTCTGTGAAGTGTGCTGCGACCCCAACCACTGCGCTTCAAGGAGGCAAA 597
    |||||
QY 1113 TGGGTTCTTGGCGCAAGCGGACCAATGTCAGG 1150
    |||||
DB 598 TGGGTTCTTGGCGCAAGCGGACCAATGTCAGG 635
    |||||

RESULT 2
US-11-075-185-35
; Sequence 35, Application US/11075185
```

Publication No. US20050266434A1
GENERAL INFORMATION:
APPLICANT: REEVES, CHRISTOPHER D
APPLICANT: JULIEN, BRYAN
APPLICANT: REID, RALPH
TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
FILE REFERENCE: 010099.03
CURRENT APPLICATION NUMBER: US/11/075,185
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/551,103
PRIOR FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: US 60/568,290
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.3
SEQ ID NO 35
LENGTH: 10968
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-11-075-185-35

Query Match 1.9%; Score 59.4; DB 7; Length 10968;
Best Local Similarity 43.6%; Pred. No. 0.0023;
Matches 325; Conservative 0; Mismatches 411; Indels 9; Gaps 1;

QY 1708 CGCCGACAGAGCCGAGGACCGCGGCGCCCTCGCCGCAACGCTGTTGTGACGCCG 1767
DB 6728 CGCACAACGCGCGAGGCGCGCGCGCGCGCTGCGCTGCGCTGCTGCTGCTGCG 6787
QY 1768 CCACAACCGCGCTGACTTCGCGCGCTCGCTATGACAGCGCAACGCACTTCGCGGCA 1827
DB 6788 GTCCGAGAGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6847
QY 1828 ACGCGGCAACGCTGCTTCTTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1887
DB 6848 ACCCGGACCGAGCGCGCTGCTGACGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6907
QY 1888 CAGGCTGACCTGCA 1947
DB 6908 CCACCGCGCGCTGCG 6967
QY 1948 GTCCCGCAGTACTGCGCGCACCAAGTCGCGCGCTGCGCGCTGCGCTGCGCGCAAC 2007
DB 6968 TTGCGGAGTTGCGCTGAGGC-----GCGCGCGCGCGCGCGCGCGCGCGCGCG 7018
QY 2008 CGCGGCG 2067
DB 7019 CG 7078
QY 2068 TGGCGCGCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2127
DB 7079 CGCGCATGCG 7138
QY 2128 TGTCCGATTCAGCTGATGAGAGCGCTCTCTGATCAAGTCAATGCACTTCACGCACT 2187
DB 7139 CTTGCGCGAGCTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7198
QY 2188 CGCGGATTTTACAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2247
DB 7199 CGCGCTCG 7258
QY 2248 AAGATTGTCCTCCGCTCAAGAGCGAGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2307
DB 7259 CTCTGAGGTGCGCGCTCTACCGCGAGTGGAGATCTGGGCGCTGCGCGCGCGCGCG 7318
QY 2308 CCAAGGACATTAGCGGCTACTATGCTTCTACGACAGTAAAGCGCGCGCGCGCGCGCG 2367
DB 7319 TCGGCGACAGCATCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7378
QY 2368 CG 2427
DB 7379 ACGCTTCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7438

QY 2428 CCCCACTCTCTCTTGCAGCCCA 2452
DB 7439 CCATGCGCTCGCTGAGGCCACCA 7463

RESULT 3
US-11-075-185-1
Sequence 1, Application US/11075185
Publication No. US20050266434A1
GENERAL INFORMATION:
APPLICANT: REEVES, CHRISTOPHER D
APPLICANT: JULIEN, BRYAN
APPLICANT: REID, RALPH
TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
FILE REFERENCE: 010099.03
CURRENT APPLICATION NUMBER: US/11/075,185
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/551,103
PRIOR FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: US 60/568,290
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 78869
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-11-075-185-1

Query Match 1.9%; Score 59.4; DB 7; Length 78869;
Best Local Similarity 43.6%; Pred. No. 0.0047;
Matches 325; Conservative 0; Mismatches 411; Indels 9; Gaps 1;

QY 1708 CGCCGACAGAGCCGAGGACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1767
DB 21995 CGCAGAGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22054
QY 1768 CCACAACCGCGCTGACTTCGCGCGCTCGCTATGACAGCGCAACGCACTTCGCGGCA 1827
DB 22055 GTCCGAGAGAGCG 22114
QY 1828 ACGCGGCAACGCTGCTTCTTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1887
DB 22115 ACCCGGACCGAGCGCGCTGCTGACGTCGCGCGCGCGCGCGCGCGCGCGCGCG 22174
QY 1888 CAGGCTGACCTGCA 1947
DB 22175 CCACGCGCGCTGCG 22234
QY 1948 GTCCCGCAGTACTGCGCGCACCAAGTCGCGCGCTGCGCGCTGCGCTGCGCGCAAC 2007
DB 22235 TTGCGGAGTTGCGCTGAGGC-----GCGCGCGCGCGCGCGCGCGCGCGCGCG 22285
QY 2008 CGCGGCG 2067
DB 22286 CGCGGCGCGAGCG 22345
QY 2068 TGGCGCGCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2127
DB 22346 CGCGCATGCG 22405
QY 2128 TGTCCGATTCAGCTGATGAGAGCGCTCTCTGATCAAGTCAATGCACTTCACGCACT 2187
DB 22406 CTCTGAGGTGCGCGCTCTACCGCGAGTGGAGATCTGGGCGCTGCGCGCGCGCGCG 22465
QY 2188 CGCGGATTTTACAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2247
DB 22466 CGCGCTCGAGGCG 22525
QY 2248 AAGATTGTCCTCCGCTCAAGAGCGAGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2307
DB 22526 CTCTGAGGTGCGCGCTCTACCGCGAGTGGAGATCTGGGCGCTGCGCGCGCGCGCG 22585

QY	2308	CCAAAGACATTAAAGGGGTACTTAATAGCTTTCACTGACAGACTTAAAGCGCCGCTGCGCCG	2367
Db	22586	TCGGCCACAGATCTGGAGAGCTGCGCCGCGCCCAACATGCGCGGTGCTGACTTGCAGC	72645
QY	2368	CGAGCCCCGCGCGCGCCGAGACCCCAACCCAGACCCCTTCAAGCTTTTCCCGAGCTCGCCT	2427
Db	22646	ACGCTCTCGCCTCTGTGTGCGCGCCCGCGGGGGGAGCTCATAGAGGCTCTCCCTCGAGGAG	22708
QY	2428	CCCCACACTCTCTCTTGGCGACCCA	2452
Db	22706	CCATGAGCTTCGTGAGAGCCACGA	22730
RESULT 4			
US-10-858-730-47			
; Sequence 47, Application US/10858730			
; Publication No. US20050255568A1			
GENERAL INFORMATION:			
APPLICANT: Bailey, Richard B.			
APPLICANT: Blomquist, Paul			
APPLICANT: Doten, Reed			
APPLICANT: Driggers, Edward M.			
APPLICANT: Madden, Kevin T.			
APPLICANT: O'Leary, Jessica			
APPLICANT: O'Toole, George			
APPLICANT: Trueheart, Joshua			
APPLICANT: Walbridge, Michael J.			
APPLICANT: Yorgey, Peter S.			
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID			
FILE OF INVENTION: PRODUCTION			
FILE REFERENCE: 14184-030001			
CURRENT APPLICATION NUMBER: US/10/858,730			
CURRENT FILING DATE: 2004-06-01			
PRIOR APPLICATION NUMBER: US 60/475,000			
PRIOR FILING DATE: 2003-05-30			
PRIOR APPLICATION NUMBER: US 60/551,860			
PRIOR FILING DATE: 2004-03-10			
NUMBER OF SEQ ID NOS: 364			
SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 47			
; LENGTH: 1290			
; TYPE: DNA			
; ORGANISM: Streptomyces coelicolor			
US-10-858-730-47			
Query Match 1.7%; Score 55.2; DB 6; Length 1290;			
Best Local Similarity 45.3%; Pred. No. 0.0084;			
Matches 327; Conservative 0; Mismatches 383; Indels 12; Gaps 3;			
QY	1502	CGGCTGTGACATGAGCGCGCTGAGACCCCTGCGCCCAAGCACTGCGCGGCTGCGAGATGCT	1561
Db	51	CGGCTCAAGAGTGGGCGGATATGACAGCAGCAGCGCCGCACTCGCGCGCGGATCGG	110
QY	1562	GCCCGGGGCGCGCTTACCGCCATGCGCGGCGCTCTTCTCTGAGAGCAAGTTGTAACAATA	1621
Db	111	GGCCCCGGTGGAGCTCGGGGGGTGCGCGTACGCGGCGCCGACAAGGTGCGGAGGGGAT	170
QY	1622	CGCCAGAGCGCGCTTTCACCCGAGGCGGGGCGCGGAGCCCGGCGCGGTACAGACCGCAG	1681
Db	171	CGACCCGGCGCTCTGTCACACCGACCGACCGCGCTCTCTCAAGCGCGGGACATGAGCT	230
QY	1682	CGTCCGCGACCAACAGGAGCTGCTGCGCGCGAGCAGGAGCCAGGACCGGAGCGGCGCT	1741
Db	231	CGTCTCGAGGCTATCGGGGGGATGAGACCCGCGGAGAGCTCATACACACCGCTTTCG	290
QY	1742	GCGCGAAGCTGCTTGTGAGCGCGGAGCAACAACCTGAGCTTTCGCGGCTTGAGCTTA	1801
Db	291	GACCGGCGCTTCGTGCTGCTCGCCCAAGGCGCTCATTCGCCAGAGCGCGCGCT	350
QY	1802	TGACACGGGCAAGGACTTTCGGGGGAGCGGCGCCAGCGCTGCTCTTAAAGCGGCGG	1861
Db	351	GACGCGCGCGCGGAGACGACGCGCAAGGAC-----CTGTACTACAGAGCGCGCT	401

Qy	1862	CGTGAAGGCGCTCCGCTGCAAGGCTGACAGGCTGACGCGCGCTGCTACTACGC	1921
Db	402	CGCGGCTGCATCCGCTG--ATCCGCGCGCTGCGCGAGTCCCTGCG--CGGCGCAAGGT	458
Qy	1922	CGACCCGCTGCGGCTGCGGCGCGCGCGAGTCCCGCGAGTACTGCGGCAACAGTGGGCTC	1981
Db	459	CAACCCGGGTGCTCGGATGTGTAAGGGAACCACTTCAATCCGAGCGCATGGACTTC	518
Qy	1982	GGTGTCCCTGTGTGGCCCAACAGCGCGCGCGCGCGCATGGCGCGCCATCC	2041
Db	519	GACCGGCGCGCGCTATCAAGAGCGCTCGACGAGCGCACCGCCCTCGGGTACGCGCAAGC	578
Qy	2042	CTACTGTGGGCGAAGAGCGCGAGGCGCTGCGCGCGCGAGCGCTGCGCGCTGCCCGCGCGC	2101
Db	579	CGACCCGACCGCGCGAGTGAAGGGCTTGCACGCGCGCAACGAAGCGCGCATCTGCGCGG	638
Qy	2102	CGCGGAGAGACGCCAAGCGCCCAAGACCTGTCGATTCAGCTGATCGAGCGCCCTCTC	2161
Db	639	GATGCGCTTCAACACCGCGGTACGCTTGCACGACGTCTACCGGAGGCGATGACCGAGT	698
Qy	2162	GATCAAGTCCATCGACTCGACGAGTCTGGGGAATTACGAGCAGGCAACGAGCGGAT	2221
Db	699	CACCGCGCGCAATTGCGCTCGCGCAAGAGATGGGCTGCACCATCAAGCTGCTCGCAT	758
Qy	2222	CT 2223	
Db	759	CT 760	

RESULT 5
 ; US-11-143-980-1
 ; Sequence 1, Application US/11143980
 ; Publication No. US20050272133A1
 ; GENERAL INFORMATION:
 ; APPLICANT: He, Min
 ; APPLICANT: Hucul, John
 ; APPLICANT: Halcil, Bradley A.
 ; APPLICANT: Wagenaar, Melissa M.
 ; APPLICANT: Graziani, Edmund
 ; APPLICANT: Summers, Mia
 ; APPLICANT: Kulowski, Kerry
 ; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
 ; FILE REFERENCE: AM-101426US
 ; CURRENT APPLICATION NUMBER: US/11/143,980
 ; CURRENT FILING DATE: 2005-06-03
 ; PRIOR APPLICATION NUMBER: US 60/664,483
 ; PRIOR FILING DATE: 2005-03-23
 ; PRIOR APPLICATION NUMBER: US 60/576,895
 ; PRIOR FILING DATE: 2004-06-03
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 1
 ; LENGTH: 116856
 ; TYPE: DNA
 ; ORGANISM: Streptomyces sp.
 ; US-11-143-980-1

Qy	Query Match	1.7%;	Score 54.8;	DB 7;	Length 116856;
Db	Best Local Similarity	47.7%;	Pred. No. 0.054;		
	Matches 228; Conservative	0;	Mismatches 242;	Indels 8;	Gaps 2;

Qy	191	CGAGTGTGCTGCTGTGCGCCAAAGAGCGCGCGCGCGCGCGCATGGCGG--GGGCCAA	2038
Db	51108	CGGGGCGCTGTACTGTGCTCTCACTGGCGCGGAGACGAGTGGGCGCTTACCGGCTTCTCAC	51167
Qy	2039	TTCCTTACTTGGGCGAGAGGCCAGAGGCTTGAGCGCGCGAGCGCTGCGCTGCGCGCGG	2098

Db 51168 GCGCTTGGCCGAGGGGTACGACACGCGGTGAGGTGCGATGCGCGCTGTCCGGG 51227
QY 2099 GCGCCCGAGAGACGCGCAAGCCCAAGG-----CTGTCCGATTTCAGTGGATGAGAC 2152
Db 51228 GCGCCGCGGGGTGATCTGCGCCACGCTAGCGCTTTCAGGCGGCGCGCTCTGCTGGAAGC 51287
QY 2153 GCGCCCTGTGATCAGTCCATGCTGACCTCCAGGAGCTCGGGGATTTCAGAGGCGCAAGC 2212
Db 51288 GCGCCACCGCGAGGCGCCCGACGCGCGATCGACCGGAAATTCGTGGCGCGCTGTGACG 51347
QY 2213 GAGCGGATCTTCGCGCGCGGACGCGCGGTGTCGAGAGTTGTCGCCGCTCAAGGCA 2272
Db 51348 CGAGGACCTTCGAGTGTGCTGCGCGCGCGCTGCGCGAGCGGCGCTGCGGAAGT 51407
QY 2273 GGTGTGCGCCGAGCGGAGCTGCGAGAACTGCGCCAGAGCATTAACGCTACTAAG 2332
Db 51408 GCTGCGCGCGCTGTCCGAGTGGCGGCGAAAGCGGTGACGTCTCACATGACTCATG 51467
QY 2333 CTCTACTGCGACAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2390
Db 51468 GCGTTACACGATCCGCTGGAAGCCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCG 51525

RESULT 6

US-11-075-185-36
; Sequence 36, Application US/11075185
; Publication No. US2005026434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
; US-11-075-185-36

Query Match 1.7%; Score 54.2; DB 7; Length 5679;
Best Local Similarity 42.8%; Pred. No. 0.024;
Matches 345; Conservative 0; Mismatches 453; Indels 9; Gaps 1;

QY 1646 GCGGCGCGCGCGCGCGCGCGGTACGAGCGCGAGCGTCCGCAACCAACGCGCTGCT 1705
Db 1356 GCGGCACTGATCATGAGAGAGCGCGCGCGCGAGCGCGCGCGCGAGAGCGCGCGC 1415
QY 1706 GTGCGCGGAG 1765
Db 1416 AGCGGCGGAG 1475
QY 1766 GCGCAACACGCGGTGAGCTTTCGCGGCTTCGCGCTTATGACGCGCAACGAGTTCGCGG 1825
Db 1476 GGGTCCGAG 1535
QY 1826 CAACGCGGACGCTGCTCTTTACGCGGCGCGGCGGTGAAGAGCGCTGCGCGAGCG 1885
Db 1536 CCAACCGGACGCTGCGGTTCGAGGTGCGCGGCGCGCGAGCGAGCGAGCGAGCGACT 1595
QY 1886 TGCAGGCTGATGCGCGCGCGCGCTGAGTACGCGCAACCGTGGGCTGGGCGCGCG 1945
Db 1596 CCGCAGCGCGCTGCGCGCTGCGCGCGCGCGAGCGAGCGAGAGAGAGAGAGAGAG 1655
QY 1946 CAGTCCCGCGAGTACTGCGGAGCAACAGTCCGGCTCGGTGCGCTGCGCGCAACAG 2005

Db 1656 CTTTGCAGATTCGCGCGCGCG-----GGCCCGGCGCGCGAGCGCGCGCGCTGAC 1706
QY 2006 GCGCGCGCGCGCGCGCGCGGATGCGCGCGCGCAATTCCTACTGTGGAGAGCGAGGG 2065
Db 1707 GCGCGCGGAGAGCGCGCGCGAGAGAGTGGGTGCTCTTACCGCGCGAGGCGAGCGCG 1766
QY 2066 CTTGCGCGGAGAGCTGCGCGCTGCGCGCGCGCGCGCGCGCGAGAGAGCGCAAGCGCAAG 2125
Db 1767 GCGCGGATGGGCGCGCGCGCTGTATGCGACCGACCGGTCTTCCGCGCGCGCTGACG 1826
QY 2126 CTGTGCAATTCAGCTGATTCGAGAGCGCGCTCTCTGATCAAGTTCATGCTTCAGCG 2185
Db 1827 GCGTGGCGGAGCTCGACCGCGCACTGACCGCGCGCGCTGTGAGCGCTCTTCGCGGA 1886
QY 2186 CTGCGGAGTTTACGAGAGCGCAAGCGGAGCGGATTCGCGCGCGCGAGCGCGCGGTG 2245
Db 1887 GCGCGGCTCGAGGCGCGCGGCTGTGACAGAGAGCGCTGGAGCTGAGCGCGCGCTGT 1946
QY 2246 CGAGATTTCGCGCGCTCAAGAGCGAGTGTGCGCGCGCGAGCTGCGAGAGAACTG 2305
Db 1947 GCTCTGAGAGTGGCGCTTACGAGAGTGGAAAGCTGGGGCTGCGCGCGCGCGCT 2006
QY 2306 GCGCAAGACATTAGCGGCTACTATGCGCTTCTACTGACAGCTAGGCGCGCGCTGCG 2365
Db 2007 GCTGCGGACAGCTCGCGGAGATGTCGCGCGCGACATGCGCGCGCTGCTGCACTTCA 2066
QY 2366 CCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2425
Db 2067 CGAGCGCTCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2126
QY 2426 CTCCCACTCTCTCTTGGCGACCA 2452
Db 2127 GCGCATGCTCTCATCGAGCGCACCA 2153

RESULT 7

US-11-075-185-2/c
; Sequence 2, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 14172
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
; US-11-075-185-2

Query Match 1.7%; Score 54.2; DB 7; Length 14172;
Best Local Similarity 45.2%; Pred. No. 0.034;
Matches 286; Conservative 0; Mismatches 338; Indels 9; Gaps 2;

QY 1519 GCTGACCCCTCGCCCAAGACTGCGCGCGCTGCGAGATGTGCCCGGCGCGGCTACG 1578
Db 2636 GCGGAGACTGCTGATGAGATGACGCGCTGACCTTCTGTCTGAGGCGCTTCTCGCG 2577
QY 1579 CCAATGCGGCTCTTCTCTGAGAGACAGTTCTGAGACACTAGCCAAAGCGCGCTTC 1638
Db 2576 GCTGCGCTTCTGCGCGCGCGCGCGCGCGCTGCGCGCTGCGCTGCGCTGAGCTTCG 2517
QY 1639 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTACGAGCGCGAGCTGCGCGCAACG 1698

Db 2516 GCCTTGTGCGCGCGCGCTGTACTTCGCGCGCGCGCGCGCGCTGCTGGGCAAGACCTGACG 2457
 Oy 1699 GAGCTGTGCG 1758
 Db 2456 GACG 2397
 Oy 1759 TGACG 1818
 Db 2396 TATGCG 2337
 Oy 1819 TCG 1878
 Db 2336 TGTGCG 2277
 Oy 1879 TCGAGGCTGCAAGCTGCACTGCG 1938
 Db 2276 ACATGACGCGCG-----ACATTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2222
 Oy 1939 GCG 1998
 Db 2221 CTGCGGCG 2162
 Oy 1999 CCAACG 2058
 Db 2161 GCGGCTGCG 2106
 Oy 2059 CCGAGGCG 2118
 Db 2105 CCGGCG 2046
 Oy 2119 CCAAGGACCTGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2151
 Db 2045 TGGAGGCG 2013

RESULT 8

US-11-075-185-36/c
 ; Sequence 36, Application US/11075185
 ; Publication No. US20050266434A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REEVES, CHRISTOPHER D
 ; APPLICANT: JULIEN, BRYAN
 ; APPLICANT: REID, RALPH
 ; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
 ; FILE REFERENCE: 010099.03
 ; CURRENT APPLICATION NUMBER: US/11/075.185
 ; PRIOR FILING DATE: 2005-03-07
 ; PRIOR APPLICATION NUMBER: US 60/551,103
 ; PRIOR FILING DATE: 2004-03-08
 ; PRIOR APPLICATION NUMBER: US 60/568,290
 ; PRIOR FILING DATE: 2004-05-04
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 36
 ; LENGTH: 5679
 ; TYPE: DNA
 ; ORGANISM: Sorangium cellulosum
 US-11-075-185-36

Query Match 1.7%; Score 53.2; DB 7; Length 5679;
 Best Local Similarity 45.8%; Pred. No. 0.04;
 Matches 335; Conservative 0; Mismatches 388; Indels 9; Gaps 4;

Db 1527 CCCTGCGCGCAAGCACTGCG 1586
 Oy 4251 CG 4192
 Db 1587 GAGCTTTTCTGCAAGCACTGCG 1646
 Oy 4191 ACAGCAAGAGCG 4132
 Db 1647 GCGG--GCGGCG 1704

Db 4131 GGGGTAATCG 4072
 Oy 1705 TGTGCG 1764
 Db 4071 CAGCG 4012
 Oy 1765 GGGGCAACCAAGCG 1824
 Db 4011 C-GCAGTGTGCG 3953
 Oy 1825 GCAACG 1882
 Db 3952 CGTGACG 3893
 Oy 1883 GAGCTGAGGCTGCACTGCG 1938
 Db 3892 GCGGCG 3833
 Oy 1939 GCG 1998
 Db 3832 GACG 3773
 Oy 1999 CCAACG 2058
 Db 3772 CCAAGCG 3713
 Oy 2059 CCGAGGCG 2118
 Db 3712 GAGCTGCG 3653
 Oy 2119 CCAAGGACCTGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2178
 Db 3652 CGGTC 3593
 Oy 2179 CCAAGGACCTGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2238
 Db 3592 CCGTC 3533
 Oy 2239 CCGTGTCCGAGA 2250
 Db 3532 CCGTCGACGCGA 3521

RESULT 9

US-11-075-185-1/c
 ; Sequence 1, Application US/11075185
 ; Publication No. US20050266434A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REEVES, CHRISTOPHER D
 ; APPLICANT: JULIEN, BRYAN
 ; APPLICANT: REID, RALPH
 ; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
 ; FILE REFERENCE: 010099.03
 ; CURRENT APPLICATION NUMBER: US/11/075.185
 ; PRIOR FILING DATE: 2005-03-07
 ; PRIOR APPLICATION NUMBER: US 60/551,103
 ; PRIOR FILING DATE: 2004-03-08
 ; PRIOR APPLICATION NUMBER: US 60/568,290
 ; PRIOR FILING DATE: 2004-05-04
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 1
 ; LENGTH: 7869
 ; TYPE: DNA
 ; ORGANISM: Sorangium cellulosum
 US-11-075-185-1

Query Match 1.7%; Score 53.2; DB 7; Length 7869;
 Best Local Similarity 45.8%; Pred. No. 0.1; 388; Indels 9; Gaps 4;
 Matches 335; Conservative 0; Mismatches 388; Indels 9; Gaps 4;

Db 1527 CCCTGCGCGCAAGCACTGCG 1586

```
Db 30482 CGCGCGCGCTTCGCGCTGCGACCGAAAGCTCCGGCTCCGCCCCGCGCGACGCGCCG 30423
QY 1587 GCGTCTTTCCTGCGAGAGACAGTTCGTGAGCAACTACCGCAAGGCCCGCTTCCACCCCGGCG 1646
Db 30422 ACGCAAGAGCGCGCGCTCGACAGGCGCTCGCGCCCGACGTCACACACCCGACCCGCGCGT 30363
QY 1647 GCGG--GCGCGGCGCCCGCGCGCGAGGTACGAGACCGAGCGTGCAGACACCAAGCGGCTGC 1704
Db 30362 GGGGACTCTGCGCGCGCGTGGCCCCGACAGAGGCCCGACAGCGCGCGCGCCCGACGCGCTC 30303
QY 1705 TGTCCCGCAGCAGCGCCGAGAGACCGCGGCGCGCGCTCGCGCGACGCTGTTTGTGACGC 1764
Db 30302 CACGCGCTTGTCTCGGCGCGCGTGCACCGCGCAGCGCGTCAACCAAGAGTCTCGTGC 30243
QY 1765 CGGCGCAACACCGCGTGAATTCCTGCGGCTCGCGCTTAAGACAGCGCGCAACGACTTCGCGG 1824
Db 30242 C-GCCAGCTGCGCGCGCGCGAGACCACTCTGACACAGGCGCGCGCGCGCGCGACCGCGG 30184
QY 1825 GCAACGCGCGCACGCTGCTCTTACGCGCGCGCGCGAGGTGAAGCGCTG--CCGCTGCA 1882
Db 30183 CGTGAACGCGCGCGCGACGACCGCGCGCTGATCGCGCGCGTCAAGTCAAGACGCGCGC 30124
QY 1883 GCGTCAAGCTGCACTGCGCGCGCGCTC---GCTACTACGCGCGACCGCTCGCGCTGCG 1938
Db 30123 GCGGCGCGCGCGCGCGCTCGCGCGCGCTCGAGAGCTCGCGCGAGCGCGCGCGCGCGC 30064
QY 1939 GCGCGCGCACTCCCGCGAGTACTGCGGACCAAGTGGGCTCGGTCGCTCGCTGCTGCG 1998
Db 30063 GCACGCCACGCGCGCGCGCGAGCGACCGCTACCGCGCGACGAGTGTCTCTGCGGCG 30004
QY 1999 CCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGG 2058
Db 30003 CCGAGCTGCGCGCGCTCGACGCGCGACGTCCTGCGACGCGCGCGCGCGCGCTCGCGC 29944
QY 2059 CCGAGGCGCTGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2118
Db 29943 GCGTCCGCGCGAGACCGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCA 29884
QY 2119 CCAAGGACCTGCTCGCAATTCAGCTGAGTGAAGCGCGCTCGCTCGATCAAGTCCATCGACT 2178
Db 29883 CGGTTCTGCTCGTGCCTGCGCGCGCGCGACAGCGACGCGCGCGCGCGCGCGCTCGC 29824
QY 2179 CCAGCGACTCGGCGATTTTACGAGCAGCGCGCGCGAGCGGATCTGCGCGCGCGCGAGC 2238
Db 29823 CCTGCTCCAGCGCGCGATGCGCGCGCGCGCGCGCGCTCGCGTCCGCTCAAGCGCGCACT 29764
QY 2239 CCGTGTCCGAGA 2250
Db 29763 CCGTCCAGCGCA 29752

RESULT 10
US-11-110-082-20
; Sequence 20, Application US/11110082
; Publication No. US2005026558A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeren
; APPLICANT: Hall, Claire
; APPLICANT: Norris, Michael Geoffrey
; APPLICANT: Saulsbury, Keith Martin
; TITLE OF INVENTION: Compositions Isolated from Forage
; FILE REFERENCE: 11000.1074UC1
; CURRENT APPLICATION NUMBER: US/11/110,082
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 60/563,723
; PRIOR FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 10/655,799
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/408,782
; PRIOR FILING DATE: 2002-09-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 20
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Festuca arundinacea
US-11-110-082-20

Query Match 1.6%; Score 52.2; DB 7; Length 2302;
Best Local Similarity 44.3%; Pred. No. 0.047;
Matches 213; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

QY 1706 GTCCCGCAGAGCGCGCGAGACCGCGGCGCGCTCGCGCGCAAGCTGTTTGTGACGCC 1765
Db 585 GCCCGCGACACTGTCTCCCGACCTGTCTGCGCACTCCCGCGGACCCCAAGCGGATGCG 644
QY 1766 GCGCAACACCGGCTGAGACTTCGCGGCTCGGCGCTTAAGACAGCGCGCAACGACTTCGCGG 1825
Db 645 TACCGGCGGCGGCGACAGCTCTCTCTCAATCTGTAATCTCTCTCGCGGCTGCGT 704
QY 1826 CAACGCGCGCACGCTGCTCTTACGCGGCGCGCGCGAGGTGAAGCGCTGCGCTGACAGC 1885
Db 705 GGTGAGAGCGCGCTCGCGCGCGCGCGCGCGAGGCGCAAGCCATCGCGCTGCGGTG 764
QY 1886 TCGAGGCTGCACTGCGCGCGCGCTGCGCTACTACCGCGACCGCTGCGGCTGCGGCGCG 1945
Db 765 GCGCGACACGCGAGAGCGAGGATCGCGCTGTGCAAGCGCTGCGCGCTGCGCGGAGCG 824
QY 1946 CAGTCCCCCGAGTACGCGGCAACAGTGGGCGTGGTGTGCGCTGCGCGCGCGCGCAAG 2005
Db 825 GTGACAGAGAGAACTTCTGCGCGCGCGCGAGCGCTGTGAAGCATACCTTGTGCGC 884
QY 2006 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGG 2065
Db 885 GCGCTTCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 944
QY 2066 CTTGCGCGCGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2125
Db 945 CCGGCTTCTCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1004
QY 2126 CCGTCCGATTCAGCTGAGTTCGAGAGCGCGCTCGTATGATGATGATGATGATGATGATGAT 2185
Db 1005 CCGTCCGAGCGCGCGCTTCAAGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1064
QY 2186 C 2186
Db 1065 C 1065

RESULT 11
US-11-045-802-25/c
; Sequence 25, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Helencjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: AP2 Domain Transcription Factor ODP2 (Ovule Development Protein 2
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; PRIOR FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
```


Accession	Sequence	Position
Db	GGCGGCGGGCGGCGGGCGGCGCAGTGGGGCCCCCGGCGAGGGAGGACACGGGAGCCGTCCG	5959
Qy	CTGGGGCGCCCGCAGTCCCCCGCAGTACTGGGGACCAAG--TGGGGTCGGTGTGCC	1991
Db	CCGCGCGCATTTGGCCCCGGGCGGCGGCGCGCCGACCAACCATTTGGCGCGCGCTGGAC	5939
Qy	TGCTGGCCCAACAGCGCGCGGCGCGCGCGCAGTGGCGGCGCCATCTTACTTGGGC	2051
Db	CATTGTTCATCTGGCGCGGCGGGCGGACCGCTGGACCCGCGAATCTTCGGA	5875
Qy	GAGGAGGCGGAGGCGCTGGCGCGGCGGAGCGGCTTCGCGCTGGCGCCCGGCGCGCGGA	2107
Db	GAAACTTGAGGAGGCGCGGCGGGAGGAACGAGAGGAGGGGCGGGCGCTGGCGCA	5819

```

RESULT 14
US-10-432-483-22
: Sequence 22, Application US/10432483
: Publication No. US20050260699A1
: GENERAL INFORMATION:
: APPLICANT: desouza, Mervyn L.
: APPLICANT: Jessen, Holly
: APPLICANT: Schroeder, William A.
: APPLICANT: Gokarn, Ravi R.
: TITLE OF INVENTION: CAROTENOID BIOSYNTHESIS
: FILE REFERENCE: 12799-0020U1
: CURRENT APPLICATION NUMBER: US/10/432,483
: CURRENT FILING DATE: 2003-11-10
: PRIOR APPLICATION NUMBER: PCT/US01/43906
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US 60/252,749
: PRIOR FILING DATE: 2000-11-22
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 22
: LENGTH: 996
: TYPE: DNA
: ORGANISM: Micrococcus luteus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(993)
: US-10-432-483-22

```

```

Query Match 1.6%; Score 51.6; DB 6; Length 996;
Best Local Similarity 51.3%; Pred. No. 0.046;
Matches 120; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 1511 CATGAGCCGCTGACCCCTTGCCCAACGACTCGCGGCTTCGAGATCGTCCCGGAGC 1570
      |||||
Db 489 CATGAGCATGCCCCGGCACCCGGCCCAACCCCGAGAGCGGGAGATGTGCGGCAC 548
      |||||

QY 1571 CCGCTACGCCATGCGCGGCTCTTTCCTGACAGACCAATTCGTGACCACTACGCCAAGC 1630
      |||||
Db 549 GGGCCGCGGGCTGGGTGCAGGCTTCAAGAGTCACTTCCGCGGATCTCGGGCGAGA 608
      |||||

QY 1631 CGGCTTCCACCCGGCGGGCGGGCGCGGCGCCGGGCTACGGACCCGAGGTGCCCCA 1690
      |||||
Db 609 CCAAGACCAAGCTCGAGACGACACTTATTTCCCGGGCGGGAGCCCTCCACCTGAGAGGAAAC 668
      |||||

QY 1691 CACCAACGGGCTGTCTGTGCGCGGACGACGAGCCTCGGCGCGGCTTGCC 1744
      |||||
Db 669 CCGCAAGCGGCTGTCTGTGCGGACCTCGGCGCGGACCTGAGCGCGGCGGTGCC 722
      |||||

RESULT 15
US-10-432-483-49
; Sequence 49, Application US/10432483
; Publication No. US20050260699A1
; GENERAL INFORMATION:
; APPLICANT: desouza, Mervyn L.
; APPLICANT: Jessem, Holly
; APPLICANT: Schroeder, William A.

```

Query	Best Local Similarity	Matches	Score	Pred.	No.	DB	Length	Gaps
1511	CATGGACCGGCTAGACCCCTTCGCCCAACGACTCGCGCGCTCGAGATCGGCCGGAGGC	1570	1.6%	51.6	6		6941	0
2276	CATGGAATGCGCCCGGCAACCCGCGCCAGAACCCCGGACAGAGGAGATGCTGCCGCCAC	2335	5.3%	0.095				0
1571	CCGCTATGCGCATGAGCGCGACTCTTTCCTGACGAGACAGATTGCTGAGCACTTACGCCAAGC	1630						
2336	GGCCCGCGCGGCTGGTGTCCCGCTTCAGAAAGGTCAATTCTGTGGGATCTCGCGCGGA	2395						
1631	CCGCTTCAACCCGGGCGCGGAGCGCGGCGCCCGGAGCGGTAACGACCCGACGTCGCCGA	1690						
2396	CCACGACACAGTCTCGAGACGCACTTACTTCCCGCGCGCGAACCCCTCCACCTGACGAGAC	2455						
1691	CACCAACGGGCGTGTCTGACCGACGACGAGCCCGGAGCCCGGCGCGGCTTCGAC	1744						
2456	CCGCAACGGGCTGTGCTCGGAGACTTCGAGGCGGACCTTGACCGGCGCGTGC	2509						

Search completed: December 20, 2005, 16:52:30
Job time : 352 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 17:22:49 ; Search time 347 Seconds
(without alignments)
4500.606 Million cell updates/sec

Title: US-10-649-400-3
Perfect score: 3186
Sequence: 1 caggtgattatccctaataa.....caccacaacaccttaag 3186

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 4161431 seqs, 245089505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8322862

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications_NA_New.*
1: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	0.8	169725	7 US-11-121-086-63	Sequence 63, Appl
2	25	0.8	169725	7 US-11-121-086-63	Sequence 168788
3	25	0.8	169725	7 US-11-121-086-63	Sequence 168790
4	25	0.8	169725	7 US-11-121-086-63	Sequence 168791
5	25	0.8	169725	7 US-11-121-086-63	Sequence 168792
6	25	0.8	169725	7 US-11-121-086-63	Sequence 168793
7	25	0.8	169725	7 US-11-121-086-63	Sequence 168794
8	25	0.8	169725	7 US-11-121-086-63	Sequence 168795
9	25	0.8	169725	7 US-11-121-086-63	Sequence 168796
10	25	0.8	169725	7 US-11-121-086-63	Sequence 168797
11	25	0.8	169725	7 US-11-121-086-63	Sequence 168798
12	25	0.8	169725	7 US-11-121-086-63	Sequence 168799
13	25	0.8	169725	7 US-11-121-086-63	Sequence 168800
14	25	0.8	169725	7 US-11-121-086-63	Sequence 168801
15	25	0.8	169725	7 US-11-121-086-63	Sequence 168802
16	25	0.8	169725	7 US-11-121-086-63	Sequence 168803
17	25	0.8	169725	7 US-11-121-086-63	Sequence 168804
18	25	0.8	169725	7 US-11-121-086-63	Sequence 168805
19	25	0.8	169725	7 US-11-121-086-63	Sequence 168806
20	25	0.8	169725	7 US-11-121-086-63	Sequence 168807
21	25	0.8	169725	7 US-11-121-086-63	Sequence 168808
22	25	0.8	169725	7 US-11-121-086-63	Sequence 168809
23	25	0.8	169725	7 US-11-121-086-63	Sequence 168810

c	24	0.8	51917	6 US-10-995-561-13338	Sequence 13338, A
c	25	0.8	110847	7 US-11-121-086-63	Sequence 11, Appl
c	26	0.8	119160	7 US-11-121-086-63	Sequence 12, Appl
c	27	0.8	126552	7 US-11-121-086-63	Sequence 1, Appl
c	28	0.8	162085	7 US-11-121-086-63	Sequence 1304, A
c	29	0.8	165156	6 US-10-995-561-13304	Sequence 61, Appl
c	30	0.8	169495	7 US-11-121-086-63	Sequence 62, Appl
c	31	0.8	200628	7 US-11-121-086-63	Sequence 6, Appl
c	32	0.8	285300	6 US-10-995-561-13236	Sequence 13236, A
c	33	0.8	305312	6 US-10-995-561-13286	Sequence 13286, A
c	34	0.8	1125000	6 US-10-995-561-13286	Sequence 13286, A
c	35	0.7	25	7 US-11-121-086-63	Sequence 23615, A
c	36	0.7	201	6 US-10-995-561-23773	Sequence 23773, A
c	37	0.7	201	6 US-10-995-561-23802	Sequence 23802, A
c	38	0.7	201	6 US-10-995-561-23805	Sequence 23805, A
c	39	0.7	201	6 US-10-995-561-30516	Sequence 30516, A
c	40	0.7	201	6 US-10-995-561-43974	Sequence 43974, A
c	41	0.7	201	6 US-10-995-561-58486	Sequence 58486, A
c	42	0.7	201	6 US-10-995-561-58499	Sequence 58499, A
c	43	0.7	201	6 US-10-995-561-58501	Sequence 58501, A
c	44	0.7	201	6 US-10-995-561-58502	Sequence 58502, A
c	45	0.7	201	6 US-10-995-561-58502	Sequence 58502, A

ALIGNMENTS

RESULT 1
US-11-121-086-63
Sequence 63, Application US/11121086
Publication No. US2005026459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121, 086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567, 570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 63
LENGTH: 169725
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-63

Query Match 0.8%; Score 26; DB 7; Length 169725;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGGATGATGATGATGATGATG 2841
DB 120068 GAGTGGATGATGATGATGATG 120093

RESULT 2
US-11-121-086-63
Sequence 168788, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma
TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121, 849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567, 949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168788


```
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168788

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2933 GTCAAGCACAAACCAAGTTCATGC 2957
DB 1 GTCAAGCACAAACCAAGTTCATGC 25

RESULT 3
US-11-121-849-168790
; Sequence 168790, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168790
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168790

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3040 AGATAATATCTTAAGAGGCTCTAGA 3064
DB 1 AGATAATATCTTAAGAGGCTCTAGA 25

RESULT 4
US-11-121-849-168791
; Sequence 168791, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168791
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168791

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 ATACTCAGTAGTAGGAGGTTTCCCA 3097
DB 1 ATACTCAGTAGTAGGAGGTTTCCCA 25
```

```
RESULT 5
US-11-121-849-168792
; Sequence 168792, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168792
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168792

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3084 TGATGGGTTTCCCACTTCTCTCA 3108
DB 1 TGATGGGTTTCCCACTTCTCTCA 25

RESULT 6
US-11-121-849-168793
; Sequence 168793, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168793
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168793

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3121 AATAATTACTATGTGCGCTTAATGCA 3145
DB 1 AATAATTACTATGTGCGCTTAATGCA 25

RESULT 7
US-11-121-849-168794
; Sequence 168794, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
```



```

; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 67304
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168794
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168794

Query Match
Best Local Similarity 100.0%; Score 25; DB 7; Length 25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3132 TGTGCTTATGACACAAATAGCT 3156
DB 1 TGTGCTTATGACACAAATAGCT 25

RESULT 8
US-10-995-561-15048/c
; Sequence 15048, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15048
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-15048

Query Match
Best Local Similarity 100.0%; Score 25; DB 6; Length 201;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2817 AGTGATGATGATGATGATGATG 2841
DB 31 AGTGATGATGATGATGATGATG 7

RESULT 9
US-10-750-185-26071
; Sequence 26071, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26071
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Bovine 19866808054269
US-10-750-185-26071
```

```

Query Match
Best Local Similarity 100.0%; Score 25; DB 6; Length 2095;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2816 GAGTGATGATGATGATGATGATG 2840
DB 344 GAGTGATGATGATGATGATGATG 368

RESULT 10
US-10-995-561-13198/c
; Sequence 13198, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13198
; LENGTH: 209822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13198

Query Match
Best Local Similarity 100.0%; Score 25; DB 6; Length 209822;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2817 AGTGATGATGATGATGATGATG 2841
DB 51465 AGTGATGATGATGATGATGATG 51441

RESULT 11
US-11-121-849-168785
; Sequence 168785, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168785
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-168785

Query Match
Best Local Similarity 100.0%; Score 24; DB 7; Length 25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2870 AAGCTGTGAAATGATTGACATA 2893
DB 2 AAGCTGTGAAATGATTGACATA 25

RESULT 12
US-10-995-561-45127
; Sequence 45127, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
```

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45127
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-45127

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGATGATGATGATGATGATG 2841
DB      137 GTGATGATGATGATGATGATG 160

RESULT 13
US-10-995-561-50737
; Sequence 50737, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50737
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50737

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGATGATGATGATGATGATG 2841
DB      101 GTGATGATGATGATGATGATG 124

RESULT 14
US-10-995-561-50739
; Sequence 50739, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50739
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50739

```

```

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGATGATGATGATGATGATG 2841
DB      100 GTGATGATGATGATGATGATG 123

RESULT 15
US-10-995-561-50759
; Sequence 50759, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50759
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50759

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGATGATGATGATGATGATG 2841
DB      99 GTGATGATGATGATGATGATG 122

Search completed: December 21, 2005, 02:00:48
Job time : 350 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 15:56:28 / Search time 15680 Seconds
(without alignments)
11549.951 Million cell updates/sec

Title: US-10-649-400-3
Sequence: 1 caggtgattatcctaataa.....cacccaacaccttaagg 3186

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size: 0

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sta:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2617	82.1	2910	6	CQ729238 Sequence
2	2259	70.9	2894	8	HSU49250 Human putat
3	1793	56.3	1793	6	CQ769652 Sequence
4	1793	56.3	1793	8	BC028289 Homo sapi
5	1694	53.2	18366	8	AC009487 Homo sapi
6	548	17.2	101040	14	AC165777 Bos lauru
7	514	16.1	514	6	AR655897 Sequence
8	384	12.1	388	6	BD024632 Sequence
9	384	12.1	388	6	AX885022 Sequence
10	380	11.9	641	8	HS338878 Homo sapi
11	367	11.5	689	8	HS3324262 Homo sapi
12	343	10.8	793	8	HS3341671 Homo sapi
13	289	9.1	289	8	AY271043 Homo sapi
14	266	8.3	693	8	HS3341518 Homo sapi
15	235	7.4	235	6	BD041913 Sequence
16	235	7.4	235	6	AX906380 Sequence
17	206	6.5	3897	9	BC058399 Mus muscu
18	206	6.5	3992	9	BC052737 Mus muscu

C	19	206	6.5	176172	14	AC127930 Rattus no
C	20	206	6.5	209914	9	AL845291 Mouse DNA
C	21	155	4.9	3814	9	U49251 Mus musculu
C	22	117	3.7	701	8	HS3341512 Homo sapi
C	23	107	3.4	671	8	HS3341522 Homo sapi
C	24	89	2.8	4570	9	AB032374 Homo sapi
C	25	84	2.6	691	8	HS3339038 Homo sapi
C	26	78	2.4	101040	14	AC165777 Bos lauru
C	27	60	1.9	60	6	CQ536907 Sequence
C	28	51	1.6	690	8	HS3325162 Homo sapi
C	29	47	1.5	784	8	HS3327080 Homo sapi
C	30	33	1.0	741	8	HS333254 Homo sapi
C	31	32	1.0	951	8	HS3325603 Homo sapi
C	32	29	0.9	174764	9	AC126257 Mus muscu
C	33	29	0.9	220887	9	AC124741 Mus muscu
C	34	29	0.9	224279	9	AC102135 Mus muscu
C	35	29	0.9	261825	9	AC146300 Mus muscu
C	36	28	0.9	188317	14	AC148957 Otlemur
C	37	28	0.9	216447	14	AC133122 Rattus no
C	38	27	0.8	60	6	BD043446 Sequence
C	39	27	0.8	60	6	AX907913 Sequence
C	40	27	0.8	532	2	AY700569 Schistoc
C	41	27	0.8	547	5	AF069393 Gallus ga
C	42	27	0.8	1246	5	AF033668 Gallus ga
C	43	27	0.8	1317	5	AY179342 Lampetra
C	44	27	0.8	2855	5	AB019787 Cynops py
C	45	27	0.8	30804	8	AC134338 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS CQ729238 2910 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15172 from Patent WO02068579.
ACCESSION CQ729238
VERSION CQ729238.1 GI:42300290

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 15172 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source 1..2910
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 82.1%; Score 2617; DB 6; Length 2910;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2907; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
279 GGGAGGTTCAGGTTCTAGAGCTAGACGCTGAGCACTGCTTCTCTTATCAG 338
1 GGGAGGTTCAGGTTCTAGAGCTAGACGCTGAGCACTGCTTCTCTTATCAG 60
QY CTTCTCAAGAAATTTCTCAGTGGACGACGCTACCACTTCAGGCGGATCCGAGCTT 398
339 CTTCTCAAGAAATTTCTCAGTGGACGACGCTACCACTTCAGGCGGATCCGAGCTT 398
DB CTTCTCAAGAAATTTCTCAGTGGACGACGCTACCACTTCAGGCGGATCCGAGCTT 120
61 CTTCTCAAGAAATTTCTCAGTGGACGACGCTACCACTTCAGGCGGATCCGAGCTT 120
QY GTTCTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAATTCACCTTG 458
399 GTTCTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAATTCACCTTG 458
DB GTTCTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAATTCACCTTG 180
121 GTTCTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAATTCACCTTG 180

QY 459 AAAAAAAAAACGAGGGGATGAAGATCAGTGAATAGACAAATTTTCTGATCCAG 518
Db 181 AAAAAAAAAATTAACGAGGGGATGACGAATCAGTGAATAGACAAATTTTCTGATCCAG 240
QY 519 GACTCACGAGGGAAGTCCAGAGAGTAATCTCTCTGTCCTTGGAAGGGGCTCTGAG 578
Db 241 GACTCACGAGGGAAGTCCAGAGAGTAATCTCTCTGTCCTTGGAAGGGGCTCTGAG 300
QY 579 CTTCGTACAGTTTGCATGAGCTCTGTCGAGATCGCTACCTCTCTCTCACTCAGCCAG 638
Db 301 CTTCGTACAGTTTGCATGAGCTCTGTCGAGATCGCTACCTCTCTCTCACTCAGCCAG 360
QY 639 CCAAGTCGCGGCGCACTGCTCCAGATGCAATGTTCCCGTACCCGCGGAGAGGAGCCG 698
Db 361 CCAAGTCGCGGCGCACTGCTCCAGATGCAATGTTCCCGTACCCGCGGAGAGGAGCCG 420
QY 699 GCGCACCCGCGCTTCTCAATCGGAGAGCCCTAGCCGCTAGATGAGCCACACCCGCTCATC 758
Db 421 GCGCACCCGCGCTTCTCAATCGGAGAGCCCTAGCCGCTAGATGAGCCACACCCGCTCATC 480
QY 759 ACCCAAGAGCTTACCAAGAGCTCTGTCGAATCTCTGCGCGGAGGATATCCCAAGGCGC 818
Db 481 ACCCAAGAGCTTACCAAGAGCTCTGTCGAATCTCTGCGCGGAGGATATCCCAAGGCGC 540
QY 819 GGGTACCCCTTACCAAGAGATAGGAGCACTCCTACCAAGAGATCCGCTTACCAAGATTC 878
Db 541 GGGTACCCCTTACCAAGAGATAGGAGCACTCCTACCAAGAGATCCGCTTACCAAGATTC 600
QY 879 TCCCTCACCGAGCCGCGGCTGAGTGCAGGCAAGCAAGATGATCCTGTCGAAGAGCCG 938
Db 601 TCCCTCACCGAGCCGCGGCTGAGTGCAGGCAAGCAAGATGATCCTGTCGAAGAGCCG 660
QY 939 CTTTGGCTGAATTTTCAAGGAGCAAGAGATGATCCTGTCGAAGAGGAGGCGC 998
Db 661 CTTTGGCTGAATTTTCAAGGAGCAAGAGATGATCCTGTCGAAGAGGAGGCGC 720
QY 999 AGTTTCTCTTTTAAAGTTTAAATTTGAGTCTGATCCCAAGGCTCATTAATAT 1058
Db 721 AGTTTCTCTTTTAAAGTTTAAATTTGAGTCTGATCCCAAGGCTCATTAATAT 780
QY 1059 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGGTT 1118
Db 781 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGGTT 840
QY 1119 CCTTGGCGGAAAGCGGACCAATGTGCAAGAAATCGGGTCTAATGCAATCGGATTC 1178
Db 841 CCTTGGCGGAAAGCGGACCAATGTGCAAGAAATCGGGTCTAATGCAATCGGATTC 900
QY 1179 CCCAACACTGGGCTCACTGAGATGGGCGCAAG-AAATCTCTTTTGGAAATTTAAACTTAC 1237
Db 901 CCCAACACTGGGCTCACTGAGATGGGCGCAAGAAATCTCTTTTGGAAATTTAAACTTAC 960
QY 1238 GAACAACAA-AGAGCTTCAATTAACATAGGAGATGAGTGTGTTTACAGTCTTGACA 1296
Db 961 GAACAACAAAGGAGAGCTTCAATTAACATAGGAGATGAGTGTGTTTACAGTCTTGACA 1020
QY 1297 AGTACAGGCGCGCTGTCATGTGTGAAAGTGAACGAGAGCGGACGAGACCTAGCC 1356
Db 1021 AGTACAGGCGCGCTGTCATGTGTGAAAGTGAACGAGAGCGGACGAGACCTAGCC 1080
QY 1357 AGCGCGCGCGCGGAGAGATTCATTTCCCTGAGCTGAGTTTCACTGCGGCTACCGGCT 1416
Db 1081 AGCGCGCGCGCGGAGAGATTCATTTCCCTGAGCTGAGTTTCACTGCGGCTACCGGCT 1140
QY 1417 ACCAGAACAGATATTTACAGATGAAAAATAGATCAACCTTTTGGAAAAAGATTTTC 1476
Db 1141 ACCAGAACAGATATTTACAGATGAAAAATAGATCAACCTTTTGGAAAAAGATTTTC 1200
QY 1477 GGGATATTTATGACAGATCTACACCGGCTGTGACATGACCGGCTGAGCCCTCGCCCA 1536
Db 1201 GGGATATTTATGACAGATCTACACCGGCTGTGACATGACCGGCTGAGCCCTCGCCCA 1260

QY 1537 ACGACTGCGCGGCTCGAGATCGTGCCCGGAGCCCGCTACGCCATGAGCGGCTCTTTC 1596
Db 1261 ACGACTGCGCGGCTCGAGATCGTGCCCGGAGCCCGCTACGCCATGAGCGGCTCTTTC 1320
QY 1597 TGCAGAACAGTTTGTGAGCACTACGCCAAAGCCCGCTTTCACATCCGAGCGCGGCGCG 1656
Db 1321 TGCAGAACAGTTTGTGAGCACTACGCCAAAGCCCGCTTTCACATCCGAGCGCGGCGCG 1380
QY 1657 GCGCGCGCGGAGTACGAGCCGAGCGTCCGACACCAAGGCGCTGTGCGCGGAGC 1716
Db 1381 GCGCGCGCGGAGTACGAGCCGAGCGTCCGACACCAAGGCGCTGTGCGCGGAGC 1440
QY 1717 AGCGCAGAGAACCCGAGCGCGCTCGCGCAACGCTGGTTGTATGACCCGCGCAACCAAC 1776
Db 1441 AGCGCAGAGAACCCGAGCGCGCTCGCGCAACGCTGGTTGTATGACCCGCGCAACCAAC 1500
QY 1777 GAGTGAATTTGCGGAGCTTGGCTATGACAGGCAAGCACTTTCGAGGCAAGCGGCGCA 1836
Db 1501 GAGTGAATTTGCGGAGCTTGGCTATGACAGGCAAGCACTTTCGAGGCAAGCGGCGCA 1560
QY 1837 GCGTGTCTCTTAAAGCGCGCGCGGAGAGGAGGCTGCGGAGGCTGAGGCTGCA 1896
Db 1561 GCGTGTCTCTTAAAGCGCGCGGAGAGGAGGCTGCGGAGGCTGAGGCTGCA 1620
QY 1897 CTGCGCGCGCGCTGAGCTACAGCGCAACCGTGGGCTGGGAGCGCGAGTCCCGCG 1956
Db 1621 CTGCGCGCGCGCTGAGCTACAGCGCAACCGTGGGCTGGGAGCGCGAGTCCCGCG 1680
QY 1957 AGTACTGCGGACCAAGTGGGCTGAGTGTGCTGCTGTGCGCAACAGCGCGCGCG 2016
Db 1681 AGTACTGCGGACCAAGTGGGCTGAGTGTGCTGCTGTGCGCAACAGCGCGCGCG 1740
QY 2017 CCGCGCGGATGCGCGGCGCAATCCCTAACCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 2076
Db 1741 CCGCGCGGATGCGCGGCGCAATCCCTAACCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
QY 2077 AGCGCTCGCGCTCG 2136
Db 1801 AGCGCTCGCGCTCG 1860
QY 2137 CCAAGTGAATGAGAGCGCTCTCTGATCAAGTCAATGATCTCAAGGAGCTCGGAGATTT 2196
Db 1861 CCAAGTGAATGAGAGCGCTCTCTGATCAAGTCAATGATCTCAAGGAGCTCGGAGATTT 1920
QY 2197 ACGAGCAGGCGCAAGCGAGGAGGAGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 2256
Db 1921 ACGAGCAGGCGCAAGCGAGGAGGAGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
QY 2257 CCGCGCTCAAGAGCGAGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2316
Db 1981 CCGCGCTCAAGAGCGAGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2040
QY 2317 TTAAGGCTACTATGAGCTTCTTACTGCAAGCTAGGCGCGCTGCGCGCGCGCGCGCG 2376
Db 2041 TTAAGGCTACTATGAGCTTCTTACTGCAAGCTAGGCGCGCGCTGCGCGCGCGCGCGCG 2100
QY 2377 CCGCGCGCGGACCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2436
Db 2101 CCGCGCGCGGACCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2160
QY 2437 CCTCTTTCGCAAGCACTCATTTTATTTGACCTCTGATGAGCGGCTGTGAGCAATTAAGT 2496
Db 2161 CCTCTTTCGCAAGCACTCATTTTATTTGACCTCTGATGAGCGGCTGTGAGCAATTAAGT 2220
QY 2497 CAGGTCTCGAGCGTATTTTAACTTTTGTGACAGGAGTCTGTGCAATTAAGTACCG 2556
Db 2221 CAGGTCTCGAGCGTATTTTAACTTTTGTGACAGGAGTCTGTGCAATTAAGTACCG 2280
QY 2557 ACCTTGAACCTTGTGTAACCTTTTGTGTTTCTTCACTTACTCTTCTGTGAGGATAT 2616
Db 2281 ACCTTGAACCTTGTGTAACCTTTTGTGTTTCTTCACTTACTCTTCTGTGAGGATAT 2340
QY 2617 CCTCTTGAATTTCCCTCCCTCGTCTTCTTCACTCTCACTTCTTCTTGTATAT 2676

OY	939	CTTTTGCGCTGAAATTTTCAACCGGCACCCMAACGGAGATGATCATACCAACAAACAGGGAGGCGC	938
Db	661	CTTTTGCGCTGAAATTTTCAACCGGCACCCMAACGGAGATGATCATACCAACAAACAGGGAGGCGC	720
OY	999	ATGTTTCCCTTTTAAATTTTAACTTTTCTGATCTCGATCCACAGGCTCATTTACATATTT	1058
Db	721	ATGTTTCCCTTTTAAATTTTAACTTTTCTGATCTCGATCCACAGGCTCATTTACATATTT	780
OY	1059	TTTGTGGATGTGATTTTGGCGGATCCCAATCATCTGGAGGTTTCAAGAGGCMAATGGGTT	1118
Db	781	TTTGTGGATGTGATTTTGGCGGATCCCAATCATCTGGAGGTTTCAAGAGGCMAATGGGTT	840
OY	1119	CCTTGCGGCAAGCGGACACCCATGTGCAAGAAATCGGGCTCTATATGATCCCGGATTTCC	1178
Db	841	CCTTGCGGCAAGCGGACACCCATGTGCAAGAAATCGGGCTCTATATGATCCCGGATTTCC	900
OY	1179	CCCAACACTGGGGCTCACTGTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACG	1238
Db	901	CCCAACACTGGGGCTCACTGTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACG	960
OY	1239	AACAACAAGAGGCTTCAATTAACAATGGGAGATGGGTGTTTACGTCCCTTGCAACAG	1298
Db	961	AACAACAAGAGGCTTCAATTAACAATGGGAGATGGGTGTTTACGTCCCTTGCAACAG	1020
OY	1299	TACCAAGCCCGCTCGCTCATGTGGTGAAATGAAACGAGACGACGACGAGACACTTAGCCAG	1358
Db	1021	TACCAAGCCCGCTCGCTCATGTGGTGAAATGAAACGAGACGACGAGACACTTAGCCAG	1080
OY	1359	CCCGGCGCGGTGACAGAGCTTCACTTTCCCTGAGACTCACTGTCATGCGCTGACACCGCTTAC	1418
Db	1081	CCCGGCGCGGTGACAGAGCTTCACTTTCCCTGAGACTCACTGTCATGCGCTGACACCGCTTAC	1140
OY	1419	CAGAACACGSAATATTAACAACAATGAAATGATCACAACCCCTTTGGCAAAAGATTTCCG	1478
Db	1141	CAGAACACGSAATATTAACAACAATGAAATGATCACAACCCCTTTGGCAAAAGATTTCCG	1200
OY	1479	GATAATTAATGACACGATCTTACACCGGCTGTGACATGACCGCTGATCCCTCGCCCAAC	1538
Db	1201	GATAATTAATGACACGATCTTACACCGGCTGTGACATGACCGCTGATCCCTCGCCCAAC	1260
OY	1539	GACTGCGCGCGCTGCGAGATCGTGTGCGGGGCGCGCTACGCGCATGAGCGCGCTTTCTCGT	1598
Db	1261	GACTGCGCGCGCTGCGAGATCGTGTGCGGGGCGCGCTACGCGCATGAGCGCGCTTTCTCGT	1320
OY	1599	CAGACACAGTTTGTGAGCAACTACGACCAAGGCGCGCTTCAACCGGCGCGGAGCGCGGCGC	1658
Db	1321	CAGACACAGTTTGTGAGCAACTACGACCAAGGCGCGCTTCAACCGGCGCGGAGCGCGGCGC	1380
OY	1659	CCCCGGGCGGGATACCGAACCGCAGCGTGTGCGACACCAACGGGCTGCTGTGCGCGCACAG	1718
Db	1381	CCCCGGGCGGGATACCGAACCGCAGCGTGTGCGACACCAACGGGCTGCTGTGCGCGCACAG	1440
OY	1719	GCGGAGGACCCGCGGCGCGCCCTCGCGGCAACGCTGTTGTGACGCGCGGACCAACACCGG	1778
Db	1441	GCGGAGGACCCGCGGCGCGCCCTCGCGGCAACGCTGTTGTGACGCGCGGACCAACACCGG	1500
OY	1779	CTGGAAGTTGCGGGCTCGGCTTATGACACCGGCACAGGACTTTCGCGGCGMAAGCGGACAG	1838
Db	1501	CTGGAAGTTGCGGGCTCGGCTTATGACACCGGCACAGGACTTTCGCGGCGMAAGCGGACAG	1560
OY	1839	CTGCACTCTTACGCGCGCGCGGCGGTGAAGGCGCTGCGCGTGAAGGCTGCAAGGCTGCACT	1898
Db	1561	CTGCACTCTTACGCGCGCGCGGCGGTGAAGGCGCTGCGCGTGAAGGCTGCAAGGCTGCACT	1620
OY	1899	GCGCGCGCGCTCGGCTTACTACGCGCAGCCGTTGCGGCTGCGGCGCGCCGACGTTCCCGCAG	1958
Db	1621	GCGCGCGCGCTCGGCTTACTACGCGCAGCCGTTGCGGCTGCGGCGCGCCGACGTTCCCGCAG	1680
OY	1959	TACTGCGGACCAAGTCCGAGCTCGGTGTGCGCTGCTGTGGCCCAACAGCGCGCGCGCGC	2018
Db	1681	TACTGCGGACCAAGTCCGAGCTCGGTGTGCGCTGCTGTGGCCCAACAGCGCGCGCGCGC	1740

QY	2019	CGCGGCATATGGCGGGCGCCCAATTCCTTACTCTGGAGGAGGCGGAGGGCCCTGGCGCGGAG	2078
Db	1741	GGGGGCATATGGCGGGCGCCCAATTCCTTACTCTGGAGGAGGCGGAGGGCCCTGGCGCGGAG	1800
QY	2079	CGCTTCGCGCGCTGGCGCGCGCGCGCGCGCGGAGGAGCGCCAAAGGCCAAGAGCATCTGTCGGATTTC	2138
Db	1801	CGCTTCGCGCGCTGGCGCGCGCGCGCGCGGAGGAGCGCCAAAGGCCAAGAGCATCTGTCGGATTTC	1860
QY	2139	AGCTGGATTCGAGACGCCCTCTCTCGATCAATGCCATCTCCAGCGACTGGGGGATTTC	2198
Db	1861	AGCTGGATTCGAGAGCGCCCTCTCTCGATCAAGCTCCATCCAGCTCCAGCGACTGGGGGATTTC	1920
QY	2199	GAGGAGGCGCAAGCGGAGGCGGATCTGCGCGCGCGCGAGACGCGCCGTCGAGAGTTCGTC	2258
Db	1921	GAGGAGGCGCAAGCGGAGGCGGATCTGCGCGCGCGCGAGACGCGCCGTCGAGAGTTCGTC	1980
QY	2259	CGCGTCAAGACGAGGCTGTGGCCAGCGGAGCTGCAGAGAACTGCGCCAAAGGACATT	2318
Db	1981	CGCGTCAAGACGAGGCTGTGGCCAGCGGAGCTGCAGAGAACTGCGCCAAAGGACATT	2040
QY	2319	AGCGGCTACTATGGCTTTACTGCGACAGCTAAGCCCGCCCTGGCGCGCGCGCGCGCG	2378
Db	2041	AGCGGCTACTATGGCTTTACTGCGACAGCTAAGCCCGCCCTTACCGCGCGCGCGCGCG	2100
QY	2379	CGGCGCGGACCGCCGAGCGCGCCCTCAAGCTTCTCCCAAGCTCGGCTCCCGCACATCC	2438
Db	2101	CGGCGCGGACCGCCGAGCGCGCCCTCAAGCTTCTCCCAAGCTCGGCTCCCGCACATCC	2160
QY	2439	TCCTTGGCGACCCACTCATTTTATTGACCCCTCGATGGCGCGCTGCGACGGAATAGTCA	2498
Db	2161	TCCTTGGCGACCCACTCATTTTATTGACCCCTCGATGGCGCGCTGCGACGGAATAGTCA	2220
QY	2499	GGTCTCCGAGGCTGATTTTAACTTTTGTGACAGCAGTCTTGCAATAGCTCACCGAC	2558
Db	2221	GGTCTCCGAGGCTGATTTTAACTTTTGTGACAGCAGTCTTGCAATAGCTCACCGAC	2280
QY	2559	CTTCAACTTTGCTGTAAACCTTTTGTTTT	2588
Db	2281	CTTCAACTTTGCTGTAAACCTTTTGTTTT	2310
RESULT 3			
CQ769652			
LOCUS			
DEFINITION Sequence 378 from Patent WO2003058021.			
ACCESSION CQ769652			
VERSION CQ769652.1			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homnidae; Homo.			
REFERENCE			
1			
Koenig-Hofman,K., Kazinski,M., Schaefer,R. and Kesper,B.			
Novel apoptosis-inducing dna sequences			
Patent: WO 2003058021-A 378 17-Jul-2003;			
JOURNAL			
Xantos Biomedicine AG (DE)			
FEATURES			
source			
1..1793			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	225	GTCTGTGATTTCTAGTATTATGATTAATAGGACTTTAAACCGAGGACGGAGGCGAG	284
Db	1	GTCTGTGATTTCTAGTATTATGATTAATAGGACTTTAAACCGAGGACGGAGGCGAG	60

285 TGTTCAGGTTCTAGAGCTATGCACTGAGACATGCTCTTCTCTTCTATCATGCTCTCC 344
 Db TGTTCAGGTTCTAGAGCTATGCACTGAGACATGCTCTTCTCTTCTATCATGCTCTCC 120
 345 AAGAAATTTCTCAAGTGTAGAGAGAGCTACCAATTCAGGCGGATTCGAGCTTCTTG 404
 Db AAGAAATTTCTCAAGTGTAGAGAGAGCTACCAATTCAGGCGGATTCGAGCTTCTTG 180
 405 CACGATCATCCCATTTATCTCGACCACTGACCACTGAGAGAGATTCACTTTGAAAAA 464
 Db CACGATCATCCCATTTATCTCGACCACTGAGAGAGATTCACTTTGAAAAA 240
 465 ATTACACAGGAGATGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
 Db ATTACACAGGAGATGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 525 CCAAGGAGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 584
 Db CCAAGGAGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 301 CCAAGGAGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 361 CACAGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 421 TCTGCGGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 481 TCTGCGGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
 705 CCGGCTCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
 Db CCGGCTCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 481 CCGGCTCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 765 GAGAGCTTCAACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
 Db GAGAGCTTCAACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 541 GAGAGCTTCAACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 825 CCGTACCCACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
 Db CCGTACCCACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 601 CCGTACCCACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 885 ACCGACCCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
 Db ACCGACCCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 661 ACCGACCCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 945 CTGAATTTTCAACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1004
 Db CTGAATTTTCAACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 721 CTGAATTTTCAACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 1005 CTTTTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 1064
 Db CTTTTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 840
 781 CTTTTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 840
 1065 GATGTGATTTTGGCGGATTCCTCAATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1124
 Db GATGTGATTTTGGCGGATTCCTCAATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 841 GATGTGATTTTGGCGGATTCCTCAATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 1125 GGCAGAGCGGACCAATGATGACCAAGTGGGTCTATATGATCCGAGTTTCCCCAC 1184
 Db GGCAGAGCGGACCAATGATGACCAAGTGGGTCTATATGATCCGAGTTTCCCCAC 960
 901 GGCAGAGCGGACCAATGATGACCAAGTGGGTCTATATGATCCGAGTTTCCCCAC 960
 1185 ACTGAGGCTCACTGATGAGCGCAAGAAATCTTTTGAATAATTTAAATTTAGACACAC 1244
 Db ACTGAGGCTCACTGATGAGCGCAAGAAATCTTTTGAATAATTTAAATTTAGACACAC 1020
 961 ACTGAGGCTCACTGATGAGCGCAAGAAATCTTTTGAATAATTTAAATTTAGACACAC 1020
 1245 AAGAGAGCTTCAATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1304
 Db AAGAGAGCTTCAATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 1021 AAGAGAGCTTCAATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 1305 CCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1364
 Db CCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 1081 CCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 1365 CGGCTGACAGAGCTTCACTTTCTGAGAGCTGATTCATGCGCGTACCGGCTTACAGAC 1424

1141 CGGCTGACAGAGCTTCACTTTCCCTGAGAGCTAGTTCAATCCCGTACACCGCTACAGAAC 1200
 1425 AGGATATTATACACACTGAAATATGATCAACCTTTTGAAGAGATTTGGGATTAAT 1484
 Db AGGATATTATACACACTGAAATATGATCAACCTTTTGAAGAGATTTGGGATTAAT 1260
 1201 AGGATATTATACACACTGAAATATGATCAACCTTTTGAAGAGATTTGGGATTAAT 1260
 1485 TATGACAGATCTACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 Db TATGACAGATCTACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1261 TATGACAGATCTACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1545 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1604
 Db CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 1321 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 1605 CAGTTCTGAGACACTACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1664
 Db CAGTTCTGAGACACTACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 1381 CAGTTCTGAGACACTACCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 1665 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1724
 1441 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 1725 GACCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
 Db GACCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 1501 GACCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 1785 TTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1844
 Db TTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 1561 TTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 1845 TTTTACGAGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 Db TTTTACGAGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1621 TTTTACGAGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1905 CCGCTGCGCTACTACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1964
 Db CCGCTGCGCTACTACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 1681 CCGCTGCGCTACTACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 1965 GGCACCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793
 Db GGCACCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793

RESULT 4
 BC029289
 LOCUS
 Homo sapiens T-box, brain, 1, mRNA (cdna clone IMAGE:5181773),
 DEFINITION
 partial cds.
 ACCESSION
 BC029289.1 GI:20379748
 VERSION
 KEYWORDS
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 1793)
 REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Useth, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Rubin, G.M., Hong, L.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y.K., S.M.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalins, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Matra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL PUBMED
12477932
2 (bases 1 to 1793)
REFERENCE
AUTHORS
CONSRTM
TITLE
JOURNAL
REMARK
COMMENT
NIH MGC Project
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guanarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H., Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNI at: <http://image.llnl.gov>
Series: IRAX Plate: 50 Row: C Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22547231.

FEATURES

source

1..1793

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5181773"

/tissue_type="Brain, Lung, Testis, adult, pooled whole"

/clone_lib="NIH MGC_115"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..>1793

/gene="TBRI"

/note="synonym: TBS-56"

/db_xref="GeneID:10716"

/db_xref="MIM:604616"

79..>1793

/gene="TBRI"

/codon_start=1

/product="TBRI protein"

/protein_id="AAH29289.1"

/db_xref="GI:62089470"

/db_xref="GeneID:10716"

/db_xref="MIM:604616"

/translation="MOLEHCLSPSTMLSKKFLNVSSYPHSGSELVLDHPITSTTD
NLERSSPLKIRITRMNTDNDNFDSKSDPDVORSKSLPVLDEVSELRHSFDSAA
DYLISQSSQPOSAATAPSAMFPYGOHPAHAFSISGSPRYMAHHVITNGAANSL
LNSSPQGYPTAGYPOQYSGOAGAPYCFSSTOPGLVPRKAOVILCNRLTKFPH
RHOETMIITKGRAMPPLSFNISGLDPIANITFVDVLADPNHMRFCGKWPVCGK
ADTNVQGNRYVYHPDSPTNGAMWROEISFGKLKLTNNKGSANNNGQNVYDOSLHRYO
PLHVEVVEDGETEDTSOPGRVQTFPFTOTIAVTAONTDIOLKIDHNEFAKGR
DNYDITGCDMDRLTPSPNDSPRSQIVGARVYAGSLDQFPVSNYAKAFHFGAG
AGPGRDTSVHTNGLISPOADBPAGSPORWFTVTPANNRLDPAASAYDTATPAG
NATLISYAAGVYKALPLQAGCTGRPLGYADPSGWARSPDYCGTSGSVLPCWP
NSAAAA"

ORIGIN

Query Match 56.3%; Score 1793; DB 8; Length 1793;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 GTCTGTGGAATTTCTAGTTTATGATAATAGCACTTTAAAAACGAGGACGGGAGCGGAC 284
DB 1 GTCTGTGGAATTTCTAGTTTATGATAATAGCACTTTAAAAACGAGGACGGGAGCGGAG 60
QY 285 TGTTCAGGTTCTAGAGCTATGACAGCTGAGACATGCTCTTCTCTTATCATGCTCTCC 344
DB 61 TGTTCAGGTTCTAGAGCTATGACAGCTGAGACATGCTCTTCTCTTATCATGCTCTCC 120
QY 345 AAGAAATTTTCATGTGAGAGAGAGCTAACCCAAATTCCAGCGGATCCGAGCTTGTCTG 404
DB 121 AAGAAATTTTCATGTGAGAGAGAGCTAACCCAAATTCCAGCGGATCCGAGCTTGTCTG 180
QY 405 CAGCATATCCCATTTATCTCGACACAGCTGACACACCTGGAGAGAGTTTCACTTTAAAAA 464
DB 181 CAGCATATCCCATTTATCTCGACACAGCTGACACACCTGGAGAGAGTTTCACTTTAAAAA 240
QY 465 ATTACAGGGGAGATGACGAATCAGTCAAGATACAGACAAATTTTCTGACTCCAGACTCA 524
DB 241 ATTACAGGGGAGATGACGAATCAGTCAAGATACAGACAAATTTTCTGACTCCAGACTCA 300
QY 525 CCAGGGGACGTCCAGAGAGTAACTCTCTCTCTTGTGACGGGGTCTGAGCTTGT 584
DB 301 CCAGGGGACGTCCAGAGAGTAACTCTCTCTCTTGTGACGGGGTCTGAGCTTGT 360
QY 585 CACAGTTTCAGTGGCTGTGCTGACAGATCGCTACCTCTCTCTCTGACGACGACAG 644
DB 361 CACAGTTTCAGTGGCTGTGCTGACAGATCGCTACCTCTCTCTCTGACGACGACAG 420
QY 645 TCTGCGGCACTGTCTCCAGTCCAGTATCCCGTATCCCGGCGGACGACGAGCGGCGAC 704
DB 421 TCTGCGGCACTGTCTCCAGTATCCCGTATCCCGGCGGACGACGAGCGGCGAC 480
QY 705 CCGGCTTCTCATTCGGGACGCTTACGCGCTACATGCGCCACACCGGCTCATACCAAC 764
DB 481 CCGGCTTCTCATTCGGGACGCTTACGCGCTACATGCGCCACACCGGCTCATACCAAC 540
QY 765 GGAGGCTTACAGCAGCTCTCTGCTCCAACTCTCGCGGAGGATACCCCAAGCGGCGTAC 824
DB 541 GGAGGCTTACAGCAGCTCTCTGCTCCAACTCTCGCGGAGGATACCCCAAGCGGCGTAC 600
QY 825 CCTTACCCACAGCAGTACGCGCACTCTTACCAAGAGCTCTGTTTACCAAGTTCTCTTC 884
DB 601 CCTTACCCACAGCAGTACGCGCACTCTTACCAAGAGCTCTGTTTACCAAGTTCTCTTC 660
QY 885 ACCCAGCGGGGCTGTGCTGCGGCAAGACAGGTGATCTGTGCAACAGGCCCCCTTGG 944
DB 661 ACCCAGCGGGGCTGTGCTGCGGCAAGACAGGTGATCTGTGCAACAGGCCCCCTTGG 720
QY 945 CTGAATTTTCAAGCGGACCAAGCGGAGTATCATCAACCAAGGAGGAGCGCATGTT 1004
DB 721 CTGAATTTTCAAGCGGACCAAGCGGAGTATCATCAACCAAGGAGGAGCGCATGTT 780
QY 1005 CTTTTTTAAGTTTAAATTTTCTGTCTGATCCCAAGCTCATTAACAATTTTTTGTG 1064
DB 781 CTTTTTTAAGTTTAAATTTTCTGTCTGATCCCAAGCTCATTAACAATTTTTTGTG 840
QY 1065 GATGTGATTTGGGGGATCCCAATCATGAGAGTTTCAAGAGGCAATGGGTCTTGTG 1124
DB 841 GATGTGATTTGGGGGATCCCAATCATGAGAGTTTCAAGAGGCAATGGGTCTTGTG 900
QY 1125 GGCAAGGAGGACCAATGTGCAAGAAATGGGTCTATATGATCCGATTTCCCAAC 1184
DB 901 GGCAAGGAGGACCAATGTGCAAGAAATGGGTCTATATGATCCGATTTCCCAAC 960
QY 1185 ACTGGGGCTCATCTGATGCGCAAGAAATCTTTTGGAAAAATTAAACTTACGAACAC 1244
DB 961 ACTGGGGCTCATCTGATGCGCAAGAAATCTTTTGGAAAAATTAAACTTACGAACAC 1020
QY 1245 AAGAGGCTTCAATTAACATAGGCGAGATGGTGTTTTACAGTCTTGCACAGTACAG 1304
DB 1021 AAGAGGCTTCAATTAACATAGGCGAGATGGTGTTTTACAGTCTTGCACAGTACAG 1080

Matches 16941, Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1493	GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAAGACTCGCGGGCTC	1552						
Db	50485	GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAAGACTCGCGGGCTC	50544						
QY	1553	GCAGATCTGTCGCGGGGCGCGCATAGCGCATAGGCGGCTCTTTCCTGACAGACAGATTGCT	1612						
Db	50545	GCAGATCTGTCGCGGGGCGCGCATAGCGCATAGGCGGCTCTTTCCTGACAGACAGATTGCT	50604						
QY	1613	GAGCAACTAGGCCAAGGCGCGCTTCCAGCCGGGCGGGGCGCGGCGCGGGTAC	1672						
Db	50605	GAGCAACTAGGCCAAGGCGCGCTTCCAGCCGGGCGGGGCGGGGCGCGGCGCGGTAC	50666						
QY	1673	GAGCCGCAAGCTGCCGACACCAAGGAGCTGTGTGCGCGACAGAGCCGAGAGACCCGGG	1732						
Db	50665	GAGCCGCAAGCTGCCGACACCAAGGAGCTGTGTGCGCGACAGAGCCGAGAGACCCGGG	50722						
QY	1733	CGGCGCCTCGCCGCAACGCTGTGTTGAGCGCCGGCAACACCGGCTGAGACTTGGCGGG	1792						
Db	50725	CGGCGCCTCGCCGCAACGCTGTGTTGAGCGCCGGCAACACCGGCTGAGACTTGGCGGG	50788						
QY	1793	CTCGGAGCTATGACACGAGCCACCGACTTTCGGGGACAGCGGACACGCTCTCTTACGC	1852						
Db	50785	CTCGGAGCTATGACACGAGCCACCGACTTTCGGGGACAGCGGACACGCTCTCTTACGC	50844						
QY	1853	GAGCGCGGAGCGTGAAGCGCTGCGGTGACAGCTGAGGCTGACTGCGCGCGCGTCCG	1912						
Db	50845	GAGCGCGGAGCGTGAAGCGCTGCGGTGACAGCTGAGGCTGACTGCGCGCGCGTCCG	50904						
QY	1913	CTACTATGCGGACCCGCTCGGAGCTGGGGCGCCGCAATCCCCCGCACTACTGGGGACCA	1972						
Db	50905	CTACTATGCGGACCCGCTCGGAGCTGGGGCGCCGCAATCCCCCGCACTACTGGGGACCA	50966						
QY	1973	GTCGGAGCTCGGCTGTGCGCTGTGGGCCCAACAGCGCGCGGCGCGCGCATGGCCGG	2032						
Db	50965	GTCGGAGCTCGGCTGTGCGCTGTGGGCCCAACAGCGCGCGGCGCGCGCATGGCCGG	51022						
QY	2033	CGCCAACTCCCTACCTTGGGCGAGGAGCGAGGGCTGGCCGCGGACGCTCGCGCTGCC	2092						
Db	51025	CGCCAACTCCCTACCTTGGGCGAGGAGCGAGGGCTGGCCGCGGACGCTCGCGCTGCC	51088						
QY	2093	GCCCCGGGCGCGCGGAGAGCGCCAAAGCCAAAGACCTGTCCGATTCCAGCTGATTCGAGAC	2152						
Db	51085	GCCCCGGGCGCGCGGAGAGCGCCAAAGCCAAAGACCTGTCCGATTCCAGCTGATTCGAGAC	51144						
QY	2153	GCCCTCCTCGATCAAGTCCATCGACTCCAGCGACTTGGGGATTTCAGAGAGGCAAGAG	2212						
Db	51145	GCCCTCCTCGATCAAGTCCATCGACTCCAGCGACTTGGGGATTTCAGAGAGGCAAGAG	51202						
QY	2213	GAGGCGGATTCGCGCGGCGACAGCGCCGTTGTCGAGAGTTGCTCCCGCTCAAGACGA	2272						
Db	51205	GAGGCGGATTCGCGCGGCGACAGCGCCGTTGTCGAGAGTTGCTCCCGCTCAAGACGA	51266						
QY	2273	GGTGCTGGCCCGAGCGGAGCTGCGAAGAACTGCGCCAAAGAGACTTATGAGGCTACTATGG	2332						
Db	51265	GGTGCTGGCCCGAGCGGAGCTGCGAAGAACTGCGCCAAAGAGACTTATGAGGCTACTATGG	51322						
QY	2333	CTTCTACTGCGACAGCTAGGCGGCGCCCTGCGCGGCGCGGCGCGCGGCGCGGACCCG	2392						
Db	51325	CTTCTACTGCGACAGCTAGGCGGCGCCCTGCGCGGCGCGGCGCGGCGCGGACCCG	51388						
QY	2393	AGCCAGGCGCTCAAGCTCTTCCCAAGCTCGGCTCCCACTCTCTCTTGCGACCGCA	2452						
Db	51385	AGCCAGGCGCTCAAGCTCTTCCCAAGCTCGGCTCCCACTCTCTCTTGCGACCGCA	51444						
QY	2453	CTCAATTTATTTGACCCCTCGATGGCGGTGCGAGCAATATGACAGGTCTCGAGCGTG	2512						
Db	51445	CTCAATTTATTTGACCCCTCGATGGCGGTGCGAGCAATATGAGGTCTCGAGCGTG	51504						
QY	2513	ATTTTAACTTTTTTGACAGAGCTCTGCGAATTAGCTACCGAGACTTCAATTGGCTG	2572						
Db	51505	ATTTTAACTTTTTTGACAGAGCTCTGCGAATTAGCTACCGAGACTTCAATTGGCTG	51566						

* 16703 16802: gap of unknown length
* 16803 16881: contig of 2079 bp in length
* 16882 18981: gap of unknown length
* 18982 22939: contig of 4018 bp in length
* 23000 23099: gap of unknown length
* 23100 25642: contig of 2543 bp in length
* 25643 25742: gap of unknown length
* 25743 31586: contig of 5844 bp in length
* 31587 31687: gap of unknown length
* 31687 34315: contig of 2629 bp in length
* 34316 37369: contig of 2954 bp in length
* 37370 39730: gap of unknown length
* 39731 39830: contig of 2261 bp in length
* 39831 42816: gap of unknown length
* 42817 42916: contig of 2986 bp in length
* 42917 48470: gap of unknown length
* 48471 48570: contig of 5554 bp in length
* 48571 51440: gap of unknown length
* 51441 51540: contig of 2870 bp in length
* 51541 53825: gap of unknown length
* 53826 53925: contig of 2285 bp in length
* 53926 56237: gap of unknown length
* 56238 56337: contig of 2312 bp in length
* 56338 59002: gap of unknown length
* 59003 59102: contig of 2665 bp in length
* 59103 67716: gap of unknown length
* 67717 67816: contig of 8614 bp in length
* 67817 71710: gap of unknown length
* 71711 74047: contig of 3894 bp in length
* 74048 74147: gap of unknown length
* 74148 77776: contig of 3629 bp in length
* 77777 77876: gap of unknown length
* 77877 86904: contig of 9028 bp in length
* 86905 87004: gap of unknown length
* 87005 89555: contig of 2551 bp in length
* 89556 89655: gap of unknown length
* 89656 91999: contig of 2344 bp in length
* 92000 92099: gap of unknown length
* 92100 95461: contig of 3362 bp in length
* 95462 95561: gap of unknown length
* 95562 97821: contig of 2260 bp in length
* 97822 97921: gap of unknown length
* 97922 101040: contig of 3119 bp in length.
Location/Qualifiers
1..101040
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-161A23"
4967..5066
/estimated_length=unknown
7582..7681
/estimated_length=unknown
11923..12022
/estimated_length=unknown
16703..16802
/estimated_length=unknown
18882..18981
/estimated_length=unknown
22000..23099
/estimated_length=unknown
25643..25742
/estimated_length=unknown
31587..31686
/estimated_length=unknown
34316..34415
/estimated_length=unknown
37370..37469
/estimated_length=unknown
39731..39830

gap /estimated_length=unknown
42817..42916
/estimated_length=unknown
gap 48471..48570
/estimated_length=unknown
gap 51441..51540
/estimated_length=unknown
gap 53826..53925
/estimated_length=unknown
gap 56238..56337
/estimated_length=unknown
gap 59003..59102
/estimated_length=unknown
gap 67717..67816
/estimated_length=unknown
gap 71711..71810
/estimated_length=unknown
gap 74048..74147
/estimated_length=unknown
gap 77777..77876
/estimated_length=unknown
gap 86905..87004
/estimated_length=unknown
gap 89556..89655
/estimated_length=unknown
gap 92000..92099
/estimated_length=unknown
gap 95462..95561
/estimated_length=unknown

Query Match 17.2%; Score 548; DB 14; length 101040;
Best Local Similarity 99.7%; Pred. No. 2.1e-307;
Matches 648; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 TAACCAATGGCAGAGCCGTTTCATCGGCTGCACACAGCAGATAAAGTGAAGCTT 110
Db 65376 TAACCAATGGCAGAGCCGTTTCATCGGCTGCACACAGCAGATAAAGTGAAGCTT 65317
QY 111 TTCTGATTTGCTGCAATGTCATTAATGGCCAAATCTCTCCAGGAGGAGGAGGAGTAA 170
Db 65316 TTCTGATTTGCTGCAATGTCATTAATGGCCAAATCTCTCTCCAGGAGGAGGAGTAA 65257
QY 171 ATCAAACTTTGAGAGCATTTCGTTGAAGTCTTTCTGTCTAGTGAAGGCTGT 230
Db 65256 ATCAAACTTTGAGAGCATTTCGTTGAAGTCTTTCTGTCTAGTGAAGGCTGT 65197
QY 231 GGATTTCTAGTTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGAGTCTTCA 290
Db 65196 GGATTTCTAGTTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGAGTCTTCA 65137
QY 291 GGTCTAGAGCTATGAGCTGAGACATGCTTCTCTCTCATGCTCTCCAGAAA 350
Db 65136 GGTCTAGAGCTATGAGCTGAGACATGCTTCTCTCTCATGCTCTCCAGAAA 65077
QY 351 TTCTCAATGTGACACAGCTACCAATTCAAGCGGATCCAGCTTGTCTTGACGAT 410
Db 65076 TTCTCAATGTGACACAGCTACCAATTCAAGCGGATCCAGCTTGTCTTGACGAT 65017
QY 411 CATCCCATTTATCTGACCACTGACCACTTGAGAGAACTTCACTTTGAAAAAATTACC 470
Db 65016 CATCCCATTTATCTGACCACTGACCACTTGAGAGAACTTCACTTTGAAAAAATTACC 64957
QY 471 AGGGGATGACGATCAGTACGATACAGACAAATTTCTGACCTCAAGAGCTCACAGGG 530
Db 64956 AGGGGATGACGATCAGTACGATACAGACAAATTTCTGACCTCAAGAGCTCACAGGG 64897
QY 531 GACGTCAGAGAACTTCTCTCTGTCTTGACGCGGCTCTGAGCTTCTGCAAGT 590
Db 64896 GACGTCAGAGAACTTCTCTCTGTCTTGACGCGGCTCTGAGCTTCTGCAAGT 64837
QY 591 TTGATGAGCTCTGCTGAGATCGTCACTCTCTCTGATGTCAGCCACCAAGTCTGCG 650
Db 64836 TTGATGAGCTCTGCTGAGATCGTCACTCTCTCTGATGTCAGCCACCAAGTCTGCG 64777

FEATURES

source
1..101040
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-161A23"
4967..5066
/estimated_length=unknown
7582..7681
/estimated_length=unknown
11923..12022
/estimated_length=unknown
16703..16802
/estimated_length=unknown
18882..18981
/estimated_length=unknown
22000..23099
/estimated_length=unknown
25643..25742
/estimated_length=unknown
31587..31686
/estimated_length=unknown
34316..34415
/estimated_length=unknown
37370..37469
/estimated_length=unknown
39731..39830

QY 651 GCCACTGCTCCAGTCCATGTTCCCGTACCCCGGCGACGACCGCG 700
 Db 64776 GCCACTGCTCCAGTCCATGTTCCCGTACCCCGGCGACGACCGCG 64727

RESULT 7
 AR655897/c 514 bp DNA linear PAT 13-JUN-2005
 LOCUS AR655897
 DEFINITION Sequence 37 from patent US 6893820.
 ACCESSION AR655897
 VERSION AR655897.1 GI:67588385
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 AUTHORS 1 (bases 1 to 514)
 TITLES
 JOURNAL Patent: US 6893820-A 37 17-MAY-2005;
 FEATURES The Ohio State University Research Foundation; Columbus, OH
 source Location/Qualifiers
 1..514
 /organism="unknown"
 /mol_type="genomic DNA"

Query Match 16.1%; Score 514; DB 6; Length 514;
 Best Local Similarity 100.0%; Pred. No. 6.3e-288; Mismatches 0; Indels 0; Gaps 0;
 Matches 514; Conservative 0;

QY 1504 GCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGCTCGACATGTGC 1563
 Db 514 GCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGCTCGACATGTGC 455
 QY 1564 CCGGGGCCGCTACGCGCATGCGCGCTCTTTCTCTGACAGACAGTTCTGAGCAATTAAG 1623
 Db 454 CCGGGGCCGCTACGCGCATGCGCGCTCTTTCTCTGACAGACAGTTCTGAGCAATTAAG 395
 QY 1624 CCAAGGCCGCTTCCACCGCGGGCGGGGGCGGGGCCCGGGCGGGGACCGGACG 1683
 Db 394 CCAAGGCCGCTTCCACCGCGGGCGGGGGCGGGGCCCGGGCGGGGACCGGACG 335
 QY 1684 TGCAGACACCAACGAGGCTGTCTGCGCGGACGAGCCGAGGACCGCGGCGCGCTCGC 1743
 Db 334 TGCAGACACCAACGAGGCTGTCTGCGCGGACGAGCCGAGGACCGCGGCGCGCTCGC 275
 QY 1744 CGCAACGCTGTTGTGACCGCGGCAACACCGGCTGACCTTCGGGCTCGGCTTAAG 1803
 Db 274 CGCAACGCTGTTGTGACCGCGGCAACACCGGCTGACCTTCGGGCTCGGCTTAAG 215
 QY 1804 ACAACGCGCAAGGACTTTCGGGGGCAAGCGGGCGACGCTCTCTTAAGCGGGCGGGGG 1863
 Db 214 ACAACGCGCAAGGACTTTCGGGGGCAAGCGGGCGACGCTCTCTTAAGCGGGCGGGGG 155
 QY 1864 TGAAGGCGCTGCGCTGACAGGCTGACATGCGCGCGCGCGCTCGGCTACTTACGCGC 1923
 Db 154 TGAAGGCGCTGCGCTGACAGGCTGACATGCGCGCGCGCGCTCGGCTACTTACGCGC 95
 QY 1924 ACCGCTGGGCTGGGGCGCGCGGAGTCCCGGAGTACTGCGGACCAAGTGGGCTCGG 1983
 Db 94 ACCGCTGGGCTGGGGCGCGCGGAGTCCCGGAGTACTGCGGACCAAGTGGGCTCGG 35
 QY 1984 TGCCTGCTGCTGGCGGCAAGCGCGCGGGCGG 2017
 Db 34 TGCCTGCTGCTGGCGGCAAGCGCGGGCGGCGC 1

RESULT 8
 BD024632 388 bp DNA linear PAT 27-AUG-2002
 LOCUS BD024632
 DEFINITION Sequence tag and encoded human protein.

ACCESSION BD024632
 VERSION BD024632.1 GI:22565855
 KEYWORDS
 SOURCE JP 2001269182-A/878.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 388)
 TITLES
 JOURNAL Edwards, U.B.D.M., Duclair, E. and Jordan, J.Y.
 GENSET Patent: JP 2001269182-A 878 02-OCT-2001;

COMMENT OS Homo sapiens (human)
 PN JP 2001269182-A/878
 PD 02-OCT-2001
 PF 24-FEB-2000 JP 2000118773
 PR 26-FEB-1999 US 60/122487
 PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
 PI JORDAN
 PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
 C12N5/10,
 PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
 G06F15/40
 CC Key Location/Qualifiers
 FH 117..386.
 FT CDS

ORIGIN
 Query Match 12.1%; Score 384; DB 6; Length 388;
 Best Local Similarity 100.0%; Pred. No. 7.3e-212; Mismatches 0; Indels 0; Gaps 0;
 Matches 384; Conservative 0;

QY 191 TTGCTGTTGAAGTCTTTCTGTCTAGTGAAGGGCTGTGATTTCTAGTTATGATTA 250
 Db 5 TTGCTGTTGAAGTCTTTCTGTCTAGTGAAGGGCTGTGATTTCTAGTTATGATTA 64
 QY 251 ATAGACTTTTAAACACGAGGACGGGAGGCGAGTTCAGGTTCTAGAGTACAGCT 310
 Db 65 ATAGACTTTTAAACACGAGGACGGGAGGCGAGTTCAGGTTCTAGAGTACAGCT 124
 QY 311 GGAGCAGTCCCTTCTCTTATCATGCTCTCCAAAGAAATTTCTCAATGTGACGACG 370
 Db 125 GGAGCAGTCCCTTCTCTTATCATGCTCTCCAAAGAAATTTCTCAATGTGACGACG 184
 QY 371 CTACCCACATTGAGCGGATCCGAGCTTGTCTGACGATCATCCCATTAATCTGACAC 430
 Db 185 CTACCCACATTGAGCGGATCCGAGCTTGTCTGACGATCATCCCATTAATCTGACAC 244
 QY 431 TGACCACTCGAGAGAGTTCACCTTGAATAAATTTACAGGGGATGACGAATCAGTC 490
 Db 245 TGACCACTCGAGAGAGTTCACCTTGAATAAATTTACAGGGGATGACGAATCAGTC 304
 QY 491 AGATACAGACAATTTTCTGACTCCAGGACTCACAGGGAGCGTCCAGAGAAATTAAT 550
 Db 305 AGATACAGACAATTTTCTGACTCCAGGACTCACAGGGAGCGTCCAGAGAAATTAAT 364
 QY 551 CTCTCTGCTTTGAGACGGGCTCTC 574
 Db 365 CTCTCTGCTTTGAGACGGGCTCTC 388

RESULT 9
 AX885022 388 bp DNA linear PAT 18-DEC-2003
 LOCUS AX885022
 DEFINITION Sequence 885 from Patent EP1033401.
 ACCESSION AX885022
 VERSION AX885022.1 GI:40040767

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Dumas Milne Edwards J.B., Duclert A. and Giordano J.Y.
AUTHORS Expressed sequence tags and encoded human proteins
TITLE Patent: EP 1033401-A 885 06-SEP-2000;
JOURNAL Genet (FR)
FEATURES
source location/Qualifiers
1..388
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
117..>386
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE99136.1"
/db_xref="GI:40040768"
/translation="MQLEHCLSPSITMSKFLNVSSSPHSGSELVHDPHISTTD
NLRSSPLKKITRGMNDSITNFPDSKDSPEDVRSKLSPLVDGV"
ORIGIN
Query Match 12.1%; Score 384; DB 6; Length 388;
Best Local Similarity 100.0%; Pred. No. 7.3e-212; Indels 0; Gaps 0;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 TTGCTGGTTGAAGTCTTTCTGTCTAGTAGAGGGGCTGTGTGATTTCTAGTTATGATTA 250
DB 5 TTGCTGGTTGAAGTCTTTCTGTCTAGTAGAGGGGCTGTGTGATTTCTAGTTATGATTA 64
QY 251 ATTAGCACTTTAAAAACAGAGGAGCGGAGCGGATTTAGAGTTCTAGAGCTATGCACT 310
DB 65 ATAGGACTTTAAAAACAGAGGAGCGGAGCGGATTTAGAGTTCTAGAGCTATGCACT 124
QY 311 GGAGCACTGCTTTCTCTCTTCTATGATGCTCTCAAGAAATTTCTCAATGATGAGCAG 370
DB 125 GGAGCACTGCTTTCTCTCTTCTATGATGCTCTCAAGAAATTTCTCAATGATGAGCAG 184
QY 371 CTACCCACATTCAGGGGATCCGAGCTGTCTTGACAGATCATCCCATTTCTCGACAC 430
DB 185 CTACCCACATTCAGGGGATCCGAGCTGTCTTGACAGATCATCCCATTTCTCGACAC 244
QY 431 TGAACAACCTGGAGAGAGTTACCTTTGAAAAAATTACAGGGGATGACGATCAGTC 490
DB 245 TGAACAACCTGGAGAGAGTTACCTTTGAAAAAATTACAGGGGATGACGATCAGTC 304
QY 491 AGATTACAGCAATTTTCTGACTCCAAAGACTCACAGGAGGAGCTCAGAGAAGTAACT 550
DB 305 AGATTACAGCAATTTTCTGACTCCAAAGACTCACAGGAGGAGCTCAGAGAAGTAACT 364
QY 551 CTCTCCTGTCTTGACGGGCTCTC 574
DB 365 CTCTCCTGTCTTGACGGGCTCTC 388
RESULT 10
HSA338878/c 641 bp DNA linear PRI 18-JUL-2002
LOCUS HSA338878/c
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
ACCESSION AJ338878
VERSION AJ338878.1 GI:15883296
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 641)
Kutsenko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M.,

TITLE
JOURNAL Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V.,
PUBMED 12136098
REFERENCE Leviaty V.G., Kolchanov N.A., Protopopov A.I., Kashuba V.I.,
AUTHORS Kiselev L.L., Wasserman W., Wahlestedt C. and Zabarovsky E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
FEATURES
source location/Qualifiers
1..641
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NL6-GN7C"
ORIGIN
Query Match 11.9%; Score 380; DB 8; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.8e-209; Indels 0; Gaps 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1638 CACCCGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG 1697
DB 380 CACCCGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG 321
QY 1698 GGGCTGTGTGCGCGCAGAGCGGAGACCCGGGCGCGCTTCTGCGCGCAACGCTGTTT 1757
DB 320 GGGCTGTGTGCGCGCAGAGCGGAGACCCGGGCGCGCTTCTGCGCGCAACGCTGTTT 261
QY 1758 GTGACGCCGGCAACAACCGGCTGGAAGTCTGCGGCGCTGCTGCTGCTGCTGCTGCTG 1817
DB 260 GTGACGCCGGCAACAACCGGCTGGAAGTCTGCGGCGCTGCTGCTGCTGCTGCTGCTG 201
QY 1818 TTGCGGGGCAACGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1877
DB 200 TTGCGGGGCAACGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 141
QY 1878 CTGACAGCTGAGGCTGACCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1937
DB 140 CTGACAGCTGAGGCTGACCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 81
QY 1938 GGGCGCGGCACTCCCGGCACTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1997
DB 80 GGGCGCGGCACTCCCGGCACTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21
QY 1998 CCAACAGCG 2017
DB 20 CCAACAGCG 1
RESULT 11
HSA324262/c 689 bp DNA linear PRI 18-JUL-2002
LOCUS HSA324262/c
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
ACCESSION AJ324262
VERSION AJ324262.1 GI:15868641
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 689)
Kutsenko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M.,
AUTHORS Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V.,
Leviaty V.G., Kolchanov N.A., Protopopov A.I., Kashuba V.I.,
Kiselev L.L., Wasserman W., Wahlestedt C. and Zabarovsky E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification

JOURNAL of the human genome
PUBMED Nucleic Acids Res. 30 (14), 3163-3170 (2002)
REFERENCE 12136098
2 (bases 1 to 689)

AUTHORS Zabarovsky E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden

FEATURES
source 1..689
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NL1-YB17C"

ORIGIN

Query Match 11.5%; Score 367; DB 8; Length 689;
Best Local Similarity 100.0%; Pred. No. 7,3e-202;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1651 GGGCGGCGCCCGGGGCGGGTACGAGACCGAGCGTGGCGACACCAACGGGCTGTCTCCG 1710
DB 367 GGGCGGCGCCCGGGGCGGGTACGAGACCGAGCGTGGCGACACCAACGGGCTGTCTCCG 308
QY 1711 CGCAGACGCGCGAGAGACCGCGGGCGCGCTCGCGCGAACGCTGGTTTGACGCGCGGCA 1770
DB 307 CGCAGACGCGCGAGAGACCGCGGGCGCGCTCGCGCGAACGCTGGTTTGACGCGCGGCA 248
QY 1771 ACAACCGGCTGCACTTCGCGGCTCGGCTTATGACACCGGCACAGGACTTCGGGGGCAAG 1830
DB 247 ACAACCGGCTGCACTTCGCGGCTCGGCTTATGACACCGGCACAGGACTTCGGGGGCAAG 188
QY 1831 CGGCGACGCGTCTCTTACGCGGGCGGGCGGTGAAGCGCTGCCGTGACAGCTGCAG 1890
DB 187 CGGCGACGCGTCTCTTACGCGGGCGGGCGGTGAAGCGCTGCCGTGACAGCTGCAG 128
QY 1891 GCTGCACTGGCGCGCGCTCGGCTGACTACGCGACCGCTCGGGGCGCGCGAGTC 1950
DB 127 GCTGCACTGGCGCGCGCTCGGCTGACTACGCGACCGCTCGGGGCGCGCGAGTC 68
QY 1951 CCGCGAGTACTCGCGGACCAAGTCGGGCTCGGCTGCTGCTTCTGCTGCCCAACAGGCGCG 2010
DB 67 CCGCGAGTACTCGCGGACCAAGTCGGGCTCGGCTGCTGCTTCTGCTGCCCAACAGGCGCG 8
QY 2011 CGGCGCG 2017
DB 7 CGGCGCG 1

RESULT 12
HSA341671 793 bp DNA linear PRI 18-JUL-2002
LOCUS HSA341671
DEFINITION Homo sapiens genomic sequence surrounding Noci site, clone
NM6-LP8R.
AJ341671
AJ341671.1 GI:15886089

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
PUBMED 12136098
REFERENCE 2 (bases 1 to 793)

AUTHORS Zabarovsky E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden

FEATURES
source 1..793
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NL6-LP8R"

ORIGIN

Query Match 10.8%; Score 343; DB 8; Length 793;
Best Local Similarity 100.0%; Pred. No. 8,3e-188;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2010 GCGCGCGCGCGCCATGCGCGCGCCATCTCTTACCTGAGCGAGAGCGCGAGGCTTG 2069
DB 1 GCGCGCGCGCGCGCCATGCGCGCGCCATCTCTTACCTGAGCGAGAGCGCGAGGCTTG 60
QY 2070 GCGCGCGCGCGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2129
DB 61 GCGCGCGCGCGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 2130 TCCGATTTCAGTGTGATGAGACCGCTCTCGATCAAGTCCATGACTCCGAGCACTCG 2189
DB 121 TCCGATTTCAGTGTGATGAGACCGCTCTCGATCAAGTCCATGACTCCGAGCACTCG 180
QY 2190 GGGATTTCAGAGCAGAGCGCAGAGCGAGATCTCGCGCGCGAGACCGCGGTGTCGAG 2249
DB 181 GGGATTTCAGAGCAGAGCGCAGAGCGAGATCTCGCGCGCGAGACCGCGGTGTCGAG 240
QY 2250 AGTTGTCCTCCGCTCAGAGCGAGGTGTCGCGCGCGAGCTGCGAGAAACTGCGCC 2309
DB 241 AGTTGTCCTCCGCTCAGAGCGAGGTGTCGCGCGCGAGCTGCGAGAAACTGCGCC 300
QY 2310 AAGGACATTAGCGGCTACTATGCGCTTCTTACTCGGACAGCTAGG 2352
DB 301 AAGGACATTAGCGGCTACTATGCGCTTCTTACTCGGACAGCTAGG 343

RESULT 13
AY271043/c 289 bp DNA linear PRI 08-JUL-2003
LOCUS AY271043
DEFINITION Homo sapiens clone SKT08-G6 putative promoter sequence.
ACCESSION AY271043
VERSION AY271043.1 GI:30721256
KEYWORDS

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL Genome Res. 13 (7), 1765-1774 (2003)
PUBMED 12805274
2 (bases 1 to 289)
AUTHORS Khambata-Ford S., Liu Y., Gleason C., Dickson M., Altman R.B.,
Batzoglou S. and Myers R.M.
TITLE Identification of promoter regions in the human genome by using a
retroviral plasmid library-based functional reporter gene assay
JOURNAL Submitted (07-APR-2003) Genetics, Stanford University School of
Medicine, 300 Pasteur Drive, Stanford, CA 94305-5120, USA

FEATURES

source

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SKT08-G6"
misc_feature <1..>289

/note="contains putative promoter region; identified by retroviral plasmid library-based functional reporter assay"

ORIGIN

Query Match	9.1%;	Score 289;	DB 8;	Length 289;
Best Local Similarity	100.0%;	Pred. No. 2.8e-156;		
Matches 289;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1671	ACGACCCGACGCTGCGCCACACCAACGAGGCTGCTGTGCGCGAGACAGGACCG	1730
QY	1671	ACGACCCGACGCTGCGCCACACCAACGAGGCTGCTGTGCGCGAGACAGGACCG <td>1730</td>	1730
Db	289	ACGACCCGACGCTGCGCCACACCAACGAGGCTGCTGTGCGCGAGACAGGACCG	230
QY	1731	GAGCGCGCCCTGCGCCGCAACGCTGTGTTTGACGCGCGGCAACACCGCTGATCTTCG	1790
Db	229	GAGCGCGCCCTGCGCCGCAACGCTGTGTTTGACGCGCGGCAACACCGCTGATCTTCG	170
QY	1791	GCCTTGGCCCTATGACACGCGCCACGGACTTTCGCGGGCAACGCGGCAACGCTGTCTTAC	1850
Db	169	GCCTTGGCCCTATGACACGCGCCACGGACTTTCGCGGGCAACGCGGCAACGCTGTCTTAC	110
QY	1851	GCGGCGGCGGCGGTGAAGCGCTGCGCTGACAGGCTGACAGGCTGACAGGCGCCGCTC	1910
Db	109	GCGGCGGCGGCGGTGAAGCGCTGCGCTGACAGGCTGACAGGCTGACAGGCGCCGCTC	50
QY	1911	GAGTACTACGCCGACCCGCTCGGAGCTGGAGGCGCGCGCACTCCCTCGCACT	1959
Db	49	GAGTACTACGCCGACCCGCTCGGAGCTGGAGGCGCGCGCACTCCCTCGCACT	1

RESULT 14				
HSA341518/c				
LOCUS	693 bp	DNA	linear	PRI 18-JUL-2002
DEFINITION	Homo sapiens genomic sequence surrounding NotI site, clone			

REFERENCE AUTHORS	TITLE
1 (bases 1 to 693) Kutsenko, A.S., Gzatalullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M., Podolski, R.M., Matushkin, Y.G., Gvanchandani, A., Muravenko, O.V., Levitsky, V.G., Kolchanov, N.A., Protopenko, A.I., Kasha, V.I., Kiselev, L.L., Wasserman, W., Wahlstedt, C. and Zharovskiy, E.R.	North flanking sequences: a tool for gene discovery and verification of the human genome
JOURNAL PUBMED REFERENCE AUTHORS TITLE	Nucleic Acids Res. 30 (14), 3163-3170 (2002) 12136098 2 (bases 1 to 693) Zhabrovskiy, E.R.
JOURNAL	Submitted (16-MAY-2001) Microbiology and Tumorigenesis Centre, Karolinska Institute, Theorells väg, 3, Box 280, Stockholm 171 77, Sweden

FEATURES	location/qualifiers
source	1..693
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/clone="NL6-LP8C"

ORIGIN

	Query Match	8.3%	Score 266	DB 8	Length 693
	Best Local Similarity	100.0%	Pred. No. 9.5e-143		
	Matches 266	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1752	TGGTTTGTCACGCGCGCCAAACACGGCGGTGACTTCGCGGCTCGGCTCATATACACGGCC	1811		
Db	266	TGGTTTGTCACCGCGCCAAACACGGCGGTGACTTCGCGGCTCGGCTCATATACACGGCC	207		

QY	1812	ACGGACTTCCGGGCAACCGGACCACGCTGTCTCTTAAACGCGGGCGGTGAAGGC	1871
Db	206	ACGGACTTCCGGGCAACCGGACCACGCTGTCTTAAACGCGGGCGGTGAAGGC	147
QY	1872	CTGGCCGCTGACGAGCTGACGAGCTGACAGCCGCCGCTCGGCTACTAAGCCGACCCGTCG	1931
Db	146	CTGGCCGCTGACGAGCTGACGAGCTGACAGCCGCCGCTCGGCTACTAAGCCGACCCGTCG	87
QY	1932	GAGTGGAGCGCCCGCAGTCCCCCGCAGTCTGCAGGACCAAGTCGGGCTCGGTCTGCC	1991
Db	86	GAGTGGAGCGCCCGCAGTCCCCCGCAGTCTGCAGGACCAAGTCGGGCTCGGTCTGCC	27
QY	1992	TGCTGGCCCAACAGCGCGCGGCGCC	2017
Db	26	TGCTGGCCCAACAGCGCGCGGCGCC	1

RESULT 15						
BD041913	BD041913	235 bp	DNA	linear	PAT 27-AUG-2002	
LOCUS						
DEFINITION	Sequence tag and encoded human protein.					
ACCESSION	BD041913					
VERSION	BD041913.1 GI:22583655					
KEYWORDS	JF 2001269182-A/18159.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					

REFERENCE
1 (bases 1 to 235)
AUTHORS
Edwards, J.B.D.M., Ducclair, E. and Jordan, J.Y
TITLE
Sequence tag and encoded human protein
JOURNAL
Patent: JP 200129182-A 18159 02-OCT-2001;

COMMENT	
OS	Homo sapiens (human)
PN	JP 2001269182-A/18159
PD	02-OCT-2001
PF	24-FEB-2000 JP 2000118773
PR	26-FEB-1999 US 60/122487
P1	JEAN BABUTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
P1	JORDAN
PC	C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
	C12N5/10,
PC	C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
	G06F15/40

FEATURES	CC	Key	Location/Qualifiers
SOURCE	FM		Location/Qualifiers
		1..235	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	

ORIGIN

Query Match	7.4%	Score 235	DB 6	Length 235
Best Local Similarity	100.0%	Pred. No. 1.1e-124		
Matches	235	Conservative 0	Mismatches 0	Indels 0
			Gaps 0	
QY	2923	TTTTATTTCGTCAAGGACACAAAACGATTCATGCTTAACCTTTTTCCTTCCTTCT		29823
DB	1	TTTTATTTCGTCAAGGACACAAAACGATTCATGCTTAACCTTTTTCCTTCCTTCT		60
QY	2983	TTGCTTTCTTCTCCTCTCATACCTTCTCTCTCTTTAAATTTCTTGAGAGA		3042
DB	61	TTGCTTTCTTCTCCTCTCATACCTTCTCTCTCTTTTAAATTTCTTGAGAGA		120
QY	3043	TAAATTTCTAAGAGCTTAGAAAATGAATATCTCAGTATGATGGTTTCCACTCT		3102
DB	121	TAAATTTCTAAGAGCTTAGAAAATGAATATCTCAGTATGATGGTTTCCACTCT		180
QY	3103	CCTCATTCGTTGCATGAATAATTAATCTATGTGCTTAATGACACAAATAGCTA		3157

Thu Dec 22 09:10:39 2005

us-10-649-400-3.olig.rge

Page 15

Db 181 CCTCAATCGTTGCATGAAATATTACTATGTGCCCTAATGCACAAATAGCTA 235

Search completed: December 20, 2005, 21:44:11
Job time : 15686 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 12:35:58 ; Search time 1805 Seconds

(without alignment) 11763.829 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186

Sequence: 1 caggtgattatcctaataa.....caccacaacaccttaagg 3186

(Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_21.*
1: geneseqn1908:*
2: geneseqn1908:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	3186	100.0	3186	12	ADL97817 Human TBR
2	1793	56.3	1793	11	ABE79130 Human apo
3	1694	53.2	6878	5	ABAI4582 Human ner
4	864	27.1	864	12	ACH87502 Human gen
5	515	16.2	515	12	ACH73791 Human gen
6	384	16.1	514	6	ABT07220 Human Cpg
7	384	12.1	388	3	AAC00887 Human sec
8	303	9.5	457	9	ACH15279 Human adu
9	281	8.8	463	13	ACF88226 Human STR
10	260	8.2	389	2	AAO39695 Expressed
11	260	8.2	389	2	AAO59107 Human bra
12	235	7.4	235	3	ACI8186 Human sec
13	60	1.9	60	6	ABN33794 Human sp1
14	32	1.0	1026	6	ABO22410 Oligonuc1
15	32	1.0	1026	6	ABO22411 Oligonuc1
16	27	0.8	60	3	AAI19701 Human sec
17	26	0.8	2439	4	AAH31915 Human olf
18	26	0.8	2439	3	AACT9868 Human sec
19	26	0.8	6140	4	AAI05394 Human rep

C	20	26	0.8	6140	4	ABL98256
C	21	26	0.8	6554	5	ABAI7685
C	22	26	0.8	8691	4	ABLI9838
C	23	26	0.8	8746	4	ABLI02984
C	24	26	0.8	12237	6	ABLI34359
C	25	26	0.8	13321	4	AAI546422
C	26	26	0.8	15576	10	ADCE6186
C	27	26	0.8	18376	4	AAK67167
C	28	26	0.8	23249	6	AAI19608
C	29	26	0.8	23885	4	AAK73173
C	30	26	0.8	23885	4	AAK70103
C	31	26	0.8	25074	13	ABD32941
C	32	26	0.8	29836	12	ADDO7012
C	33	26	0.8	32249	5	ABAI7155
C	34	26	0.8	38575	13	ACN37227
C	35	26	0.8	57243	11	ACN44826
C	36	26	0.8	90120	12	ADQ97413
C	37	26	0.8	110000	8	ABD53224_2
C	38	26	0.8	110000	13	ABD32911_2
C	39	26	0.8	110000	13	ADS99457_2
C	40	26	0.8	138941	8	ACC79695
C	41	26	0.8	143412	11	ACN44512
C	42	26	0.8	243999	8	ABE80229
C	43	26	0.8	260160	12	ADQ20017
C	44	26	0.8	263852	13	ADS99460
C	45	26	0.8	304905	11	ADP75180

ALIGNMENTS

RESULT 1	ADL97817	standard; DNA, 3186 BP.
ID	ADL97817;	
XX	ADL97817;	
DT	03-JUN-2004	(first entry)
XX		
DE	Human TBR1 DNA SeqID 3.	
XX		
KW	human; gene; ds; neuropsychiatric disorder; bipolar disorder;	
KW	schizophrenia; alpha-type II calcium/calmodulin dependent protein kinase;	
KW	CMKII-alpha; TBR1; Brachyury; neuroleptic; antidepressant; gene therapy;	
KW	mental disorder; mood disorder; psychosis; major depression.	
XX		
OS	Homo sapiens.	
XX		
FX	Key	Location/Qualifiers
FT	CDS	303..2351
FT		/*tag= a
FT		/product= "TBR1 protein"
XX		
PN	MO2004020455-A2.	
XX		
PD	11-MAR-2004.	
XX		
PF	26-AUG-2003; 2003WO-US026873.	
XX		
PR	28-AUG-2002; 2002US-0406879P.	
XX		
PR	27-FEB-2003; 2003US-0451306P.	
XX		
PA	(STRD) UNIV IELAND STANFORD JUNIOR.	
XX		
PI	Burney WR, Jones EG, Molnar M;	
XX		
DR	WPI: 2004-239152/22.	
XX		
DR	P-PDB; ADL97818.	
XX		
PT	Determining whether a subject has or is predisposed to a mental disorder,	
XX	comprises detecting the level of reagent that selectively associates with	
PT	a specified polynucleotide or polypeptide in a sample from the subject.	
XX		

Db 781 CCTTTTAACTTTTAAACATTCTGGTCTCGATCCACGGCTCATTAACAATATTTTGTG 840

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

XX Human nervous system related polynucleotide SEQ ID NO 6913.
DE Human; neurotrophic; cytostatic; dermatological; virocidic;
XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antistixling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebrotrophic; antiinflammatory;
KW antiallergic; antidiabetic; antituber; anticonvulsant; antitungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX WO200159063-A2.
XX 16-AUG-2001.
PD 17-JAN-2001; 2001WO-US001334.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232357P.
PR 14-SEP-2000; 2000US-0232358P.

PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251868P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
PS Disclosure; SEQ ID NO 6913; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/patdb/published_pcr_sequences
XX
XX
SQ Sequence 6878 BP; 2032 A; 1487 C; 1634 G; 1725 T; 0 U; 0 Other;
Query Match 53.2%; Score 1694; DB 5; Length 6878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1493 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTCGCCCAAGACTGCGCGCTC 1552
DB 2504 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTCGCCCAAGACTGCGCGCTC 2445
QY 1553 GCAGATCGTGGCCCGGCGCGCTGACCGCTGACCGCTGCTTTCTGACAGACCAAGTTCGT 1612
DB 2444 GCAGATCGTGGCCCGGCGCGCTGACCGCTGACCGCTGCTTTCTGACAGACCAAGTTCGT 2385
QY 1613 GAGCACTACGACCAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1672
DB 2384 GAGCACTACGACCAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2325
QY 1673 GAGCGCAGCGTGGCGCGCAACCAAGCGCGCTGCTGCTGCGCGCAGCAGCGCGCGCGCG 1732
DB 2324 GAGCGCAGCGTGGCGCGCAACCAAGCGCGCTGCTGCTGCGCGCAGCAGCGCGCGCGCG 2265
QY 1733 CGCGCGCTGCGCGCAACCGCTGCTTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1792
DB 2264 CGCGCGCTGCGCGCAACCGCTGCTTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2205
QY 1793 CTGCGGCTATGACACGCGCGCAAGACTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1852
DB 2204 CTGCGGCTATGACACGCGCGCAAGACTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2145
QY 1853 GCGCGGCG 1912
DB 2144 GCGCGGCG 2085

QY 1913 CTACTACGCGGACCGCGTGGCG 1972
DB 2084 CTACTACGCGGACCGCGTGGCG 2025
QY 1973 GTCGGGCTGCGTGGCG 2032
DB 2024 GTCGGGCTGCGTGGCG 1965
QY 2033 GCGCAATCCCTAATCTGCG 2092
DB 1964 GCGCAATCCCTAATCTGCG 1905
QY 2093 GCGCGGCG 2152
DB 1904 GCGCGGCG 1845
QY 2153 GCGCTCTCGATCAAGTCATCGACTCGAGCGACTCGGCGGATTTAGACGAGCGCAAGCG 2212
DB 1844 GCGCTCTCGATCAAGTCATCGACTCGAGCGACTCGGCGGATTTAGACGAGCGCAAGCG 1785
QY 2213 GAGCGGATCTGCG 2272
DB 1784 GAGCGGATCTGCG 1725
QY 2273 GGTGCTGGCCCG 2332
DB 1724 GGTGCTGGCCCG 1665
QY 2333 CTCTACTCGACAGCTAAGCG 2392
DB 1664 CTCTACTCGACAGCTAAGCG 1605
QY 2393 AGCGAGCG 2452
DB 1604 AGCGAGCG 1545
QY 2453 CTCAATTTATTTGACCGCTCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2512
DB 1544 CTCAATTTATTTGACCGCTCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1485
QY 2513 ATTTTAACCTTTTGTGACAGCAGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2572
DB 1484 ATTTTAACCTTTTGTGACAGCAGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 2573 TAAACCTTTGCTTTCTCTACTTACTCTTCTGCTGAGATTAATCTTCAATTCGCC 2632
DB 1424 TAAACCTTTGCTTTCTCTACTTACTCTTCTGCTGAGATTAATCTTCAATTCGCC 1365
QY 2633 TCCCGCTGCTTTCTCTTACTCTGCTTCTCTTCTGCTGATGAATCTTTCACCTTT 2692
DB 1364 TCCCGCTGCTTTCTCTTACTCTGCTTCTCTTCTGCTGATGAATCTTTCACCTTT 1305
QY 2693 AGGAGACCTGGGCACTGCTGTCAGGACAGCAGATTCGACCGCGCGCGCGCGCGCGCGCG 2752
DB 1304 AGGAGACCTGGGCACTGCTGTCAGGACAGCAGATTCGACCGCGCGCGCGCGCGCGCGCG 1245
QY 2753 ACATTAACCATGAGATGTTGACTGAAACCTGGAACCCAGCGCGCTCTTCTTATC 2812
DB 1244 ACATTAACCATGAGATGTTGACTGAAACCTGGAACCCAGCGCGCTCTTCTTATC 1185
QY 2813 CCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2872
DB 1184 CCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
QY 2873 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2932
DB 1124 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
QY 2933 GTCAAGGACAAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2992
DB 1064 GTCAAGGACAAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005

08-AUG-2002.
31-JAN-2002; 2002MO-US003077.
31-JAN-2001; 2001US-00775398.
(OHIS) UNIV OHIO STATE RES FOUND.
Plass C;
WPI; 2002-627436/67.
Identifying methylated CpG islands in malignant cells, useful for
diagnosing or classifying cancer, comprises the use of infrequent
cleaving, methylation-sensitive restriction enzymes and gel
electrophoresis.
Claim 17; Page 86; 13pp; English.

Db	34	TGTCCTGCTGGCCCAACAGCGCGGGCGC	1
RESULT 7			
XX	AAC00887		
ID	AAC00887	standard; cDNA; 388 BP.	
XX			
XX	AAC00887;		
AC			
XX			
DT	06-OCT-2000	(first entry)	
XX			
DE	Human secreted protein 5' EST, SEQ ID NO: 885.		
XX			
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;		
KM	gene therapy; chromosome mapping; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1033401-A2.		
XX			
XX	06-SEP-2000.		
PD			
XX			
PF	21-FEB-2000; 2000EP-00200610.		
XX			
XX	26-FEB-1999; 99US-0122487P.		
PR			
XX			
PA	(GEST) GENSET.		
XX			
PI	Dumas Milne Edwards J, Duclert A, Giordano J;		
XX			
DR	WPI; 2000-500381/45.		
XX	P-PBDB; AAC00881.		
XX			
PT	New Nucleic acid that is a 5' expressed sequence tag (5' EST) for		

```
DB 185 CTACCACTTCAGCGGATCCGAGCTTGTCTGTGACGATCATCCATTATCTGACCAAC 244
OY 431 TGACAACTCGAGAGAGATTCACTTTGAAAAAATTAACGAGGGGATGACGAATCAGTC 490
DB 245 TGACAACTCGAGAGAGATTCACTTTGAAAAAATTAACGAGGGGATGACGAATCAGTC 304
OY 491 AGATCAGACAAATTTTCTGACTCGAAGACTCACAGGGGAGCTCCAGAGAAATTAAC 550
DB 305 AGATCAGACAAATTTTCTGACTCGAAGACTCACAGGGGAGCTCCAGAGAAATTAAC 364
OY 551 CTCTCTGTCTTGAAGAGGCTCTC 574
DB 365 CTCTCTGTCTTGAAGAGGCTCTC 388

RESULT 8
ACH15279
ID ACH15279 standard; cDNA; 457 BP.
XX
AC ACH15279;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult brain cDNA #2491.
XX
KM Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 2491; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics or for physical mapping of the human genome,
CC identifying expressed genes or for forensic mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
```

```
XX
SQ Sequence 457 BP; 81 A; 168 C; 85 G; 123 T; 0 U; 0 Other;
Query Match 9.5%; Score 303; DB 9; Length 457;
Best Local Similarity 99.7%; Pred. No. 9.1e-136;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2372 CCGCGCGGGGCGGAGCCCGGAGCCCGTCACTCTTCCGAGCTCCGCTCC 2431
DB 104 CCGCGCGGGGCGGAGCCCGGAGCCCGTCACTCTTCCGAGCTCCGCTCC 163
OY 2432 ACATCTCTCTTGGGCAACCACTCATTTTAATTTGACCTCGATGGCGGTGACGGAAT 2491
DB 164 ACATCTCTCTTGGGCAACCACTCATTTTAATTTGACCTCGATGGCGGTGACGGAAT 223
OY 2492 AAGTGCAGGTCTCCGAGCGGATTTTAACCTTTTGGACAGAGTCTTCGCAATTAGCT 2551
DB 224 AAGTGCAGGTCTCCGAGCGGATTTTAACCTTTTGGACAGAGTCTTCGCAATTAGCT 283
OY 2552 CACGACCTTCAACTTTGTGTAAACCTTTGTTTGCTTACTACTCTTCTGTGGA 2611
DB 284 CACGACCTTCAACTTTGTGTAAACCTTTGTTTGCTTACTACTCTTCTGTGGA 343
OY 2612 GTTATCTCTCTACAATTCCCTCCCTCGTCTTCTTCACTCTTCTTCTT 2671
DB 344 GTTATCTCTCTACAATTCCCTCCCTCGTCTTCTTCACTCTTCTTCTT 403
OY 2672 GTTATGAACCTTTCACCTTTTGAAGACCTGGGAGTCTGTCAAGGACGCG 2725
DB 404 GGAATGAACCTTTCACCTTTTGAAGACCTGGGAGTCTGTCAAGGACGCG 457

RESULT 9
ACF88226
ID ACF88226 standard; DNA; 463 BP.
XX
AC ACF88226;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human SIRS/sepsis diagnostic marker DNA fragment 7086.
XX
KM Systemic inflammatory response syndrome; SIRS; antibacterial;
KW immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS Homo sapiens.
XX
PN WO2004087949-A2.
XX
PD 14-OCT-2004.
XX
PF 31-MAR-2004; 2004WO-EP003419.
XX
PR 02-APR-2003; 2003DE-01015031.
PR 08-AUG-2003; 2003DE-01036511.
PR 02-SEP-2003; 2003DE-01040395.
XX
PA (SIRS-) SIRS LAB GMBH.
XX
PI Russwurm S, Reinhart X, Saluz H, Straube E, Zipfel PF, Designer H;
XX
DR WPI; 2004-748070/73.
XX
PT In vitro detection of systemic inflammatory response syndrome and related
PT conditions, for e.g. monitoring progression, comprises detecting abnormal
PT expression of disease-related genes.
XX
PS Disclosure; Page: 75pp; German.
XX
CC The invention relates to a novel method for in vitro detection of
CC systemic inflammatory response syndrome (SIRS). The method comprises
CC detecting abnormal expression of disease-related genes, or their
CC associated peptides. The method of the invention demonstrates
```

CC antibacterial, immunosuppressive and antiinflammatory applications and
CC may be used for early differential diagnosis, monitoring progression,
CC assessing risk, assessing the likely response to treatment and for post
CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and
CC sepsis-like conditions. The recombinant or synthetic nucleic acid
CC sequences of the invention, or derived proteins or peptides, may be
CC useful as calibrants in assays for the specified diseases, for evaluating
CC activity or toxicity in screening for active agents and/or for
CC preparation of agents for treatment or prevention of the specified
CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic
CC marker DNA fragment of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC disclosed within the specification, however, these have not been taken
CC into account during indexing due to inconsistencies in application and
CC format

XX Sequence 463 BP; 86 A; 139 C; 97 G; 137 T; 0 U; 4 Other;

XX SQ Query Match 8.8%; Score 281; DB 13; Length 463;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-125;

Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2415 CCCAGCTCCGCTCCGACACTCTCTTGGCCACCCTCAATTTATTTGACCTCGAT 2474

DB 1 CCCAGCTCCGCTCCGACACTCTCTTGGCCACCCTCAATTTATTTGACCTCGAT 60

QY 2475 GGGCGCTGCGACGAAATAGTGAGTCCGAGCGTGATTTTAACTTTTGGACAGC 2534

DB 61 GGGCGCTGCGACGAAATAGTGAGTCCGAGCGTGATTTTAACTTTTGGACAGC 120

QY 2535 AGCTCTGCAATTAAGTCAACGACCTTCACTTGTCTGTAACCTTTTGCTTCTACT 2594

DB 121 AGCTCTGCAATTAAGTCAACGACCTTCACTTGTCTGTAACCTTTTGCTTCTACT 180

QY 2595 TACTCTTCTTGTGAGTTATCTCTTAACAATTCCTCCCTGCTTCTTCTTAC 2654

DB 181 TACTCTTCTTGTGAGTTATCTCTTAACAATTCCTCCCTGCTTCTTCTTAC 240

QY 2655 TCTTACTTCTTCTTCTTCTTGAATGAACCTTCACTTTAGG 2695

DB 241 TCTTACTTCTTCTTCTTCTTGAATGAACCTTCACTTTAGG 281

RESULT 10

AAQ39695
ID AAQ39695 standard; DNA; 389 BP.

XX AAQ39695;

XX 25-MAR-2003 (revised)

XX 20-MAY-1993 (first entry)

XX Expressed Sequence Tag human gene marker EST00360.

XX expressed sequence tag; human genome project; chromosome;

XX human gene sequencing; PCR mapping; somatic cell hybrids;

XX sublocalisation; gene tagging; tissue typing.

XX Synthetic.

XX WO9300353-A1.

XX 07-JAN-1993.

XX 19-JUN-1992; 92WO-US005222.

XX 20-JUN-1991; 91US-00716831.

XX 12-FEB-1992; 92US-00837195.

XX (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX

PI Venter JC, Adams MD;

XX WPI, 1993-036325/04.

XX Particular expressed sequence tags from human CDNA - corresponds to

XX transcription prods. of genes, useful for tagging genes, mapping

XX chromosomes and tissue typing.

XX Claim 3; Page 105; 199pp; English.

XX This sequence represents an EST (expressed sequence tag) ESTs are markers

XX for human genes actually transcribed in vivo. Unlike the random genomic

XX DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.

XX The use of ESTs could facilitate the tagging of most expressed human

XX genes within a few years at a fraction of the cost of complete genomic

XX sequencing. Using PCR primers AAQ39419-039580 (sequences designed from

XX the ESTs) sublocalisation of an EST can be achieved with panels of

XX fragments from specific chromosomes or pools of large genomic clones in

XX an analogous manner. This sequence represents EST00360. (Updated on 25-

XX MAR-2003 to correct PN field.)

XX SQ Query Match 8.2%; Score 260; DB 2; Length 389;
XX Best Local Similarity 99.7%; Pred. No. 6.3e-115;

Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2775 TCTAGAACCTGACCCACCCAGCGGCTCTTCTTATCCCGAGTGATGATGATGGA 2834

DB 1 TCTAGAACCTGACCCACCCAGCGGCTCTTCTTATCCCGAGTGATGATGATGGA 60

QY 2835 TGGATGTAAGGAGCTTAATTAATTTAGTGAACAAGCCTGTGAATGATTTGATAG 2894

DB 61 TGGATGTAAGGAGCTTAATTAATTTAGTGAACAAGCCTGTGAATGATTTGATAG 120

QY 2895 TGTTAATTTATTTAGTGAAGTCTAGTTTATTTCTGTCGAAGCACAACCAAGTTCA 2954

DB 121 TGTTAATTTATTTAGTGAAGTCTAGTTTATTTCTGTCGAAGCACAACCAAGTTCA 180

QY 2955 TGTCTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 3014

DB 181 TGTCTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 240

QY 3015 TTCTCTCTTTTAAATTTCTTGTGAAGTAAATATTTCTAAGGCGCTGAACATGAAT 3074

DB 241 TTCTCTCTTTTAAATTTCTTGTGAAGTAAATATTTCTAAGGCGCTGAACATGAAT 300

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

QY 3075 ACTCAGTAGTG 3085

DB 301 ACTCAGTAGTG 311

```
XX 12-FEB-1992; 92US-00837195.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX Venter CJ, Adams MD, Moreno RF;
XX WPI: 1993-272882/34.
XX Enriched oligonucleotides and corresp. sequences - used as markers for
PT human genes transcribed in-vivo, facilitate tagging of most human genes.
XX
XX Example 1; Page 122; 500bp; English.
XX
XX The Expressed Sequence Tag was isolated from a human brain cDNA library
CC as part of a large set of ESTs which can be used as markers for human
CC genes transcribed in vivo. They can be used to facilitate tagging of most
CC human genes, for mapping locations of expressed genes on chromosomes, for
CC individual or forensic identification, for mapping locations of disease-
CC associated genes, for identification of tissue type, and for prepn. of
CC antisense sequences, probes and constructs. EST00360 has a "poor" coding
CC probability as evaluated using the coding-region prediction program CRM.
CC See also AA059041-Q61440. (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 389 BP; 97 A; 81 C; 71 G; 139 T; 0 U; 1 Other;
Query Match 8.2%; Score 260; DB 2; Length 389;
Best Local Similarity 99.7%; Pred. No. 6.3e-115;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2775 TCTAGAACCTGGACCCACCACGCGCTCTTCTTATCCCGAGTGATGATGATGGA 2834
DB 1 TCTAGAACCTGGACCCACCACGCGCTCTTCTTATCCCGAGTGATGATGATGGA 60
QY 2835 TGGATGTGAGGATGTTAATATTTAGTGAACAAAGCTGTGAATGATGTCATAG 2894
DB 61 TGGATGTGAGGATGTTAATATTTAGTGAACAAAGCTGTGAATGATGTCATAG 120
QY 2895 TGTTAATTTATGTGAAGATGCTAGTTTATCTCGTCAAGGACAAACCAAGTTCA 2954
DB 121 TGTTAATTTATGTGAAGATGCTAGTTTATCTCGTCAAGGACAAACCAAGTTCA 180
QY 2955 TGCCTTAACCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTC 3014
DB 181 TGCCTTAACCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTC 240
QY 3015 TTCTCTCTCTTTAATTTCTTGAGATATATTTCTAAGAGCTCTAAGAAACATGAAT 3074
DB 241 TTCTCTCTCTTTAATTTCTTGAGATATATTTCTAAGAGCTCTAAGAAACATGAAT 300
QY 3075 ACTCAGTAGTG 3085
DB 301 ACTCAGTAGTG 311
RESULT 12
AAC18168
ID AAC18168 standard; cDNA; 235 BP.
XX AAC18168;
AC
XX
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 22243.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
PN
XX
XX 06-SEP-2000.
PD
```

```
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR (GENSET ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 22243; 71bp + Sequence Listing; English.
PS
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
SQ Sequence 235 BP; 56 A; 57 C; 27 G; 95 T; 0 U; 0 Other;
Query Match 7.4%; Score 235; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2293 TTTATTTCTGCTGAAGGACAAACCAAGTTCATGCTTACCTTTTTCCTTCTTCT 2292
DB 1 TTTATTTCTGCTGAAGGACAAACCAAGTTCATGCTTACCTTTTTCCTTCTTCT 60
QY 2293 TTVGCTTTTCCTTCTCTCTCATACTTCTCTTCTCTTCTTCTTCTTCTTCTTCT 3042
DB 61 TTVGCTTTTCCTTCTCTCTCATACTTCTCTCTCTTCTTCTTCTTCTTCTTCTGAGA 120
QY 3043 TTAATTTCTAAGAGGCTCTAAGAAACATGAATATCTCAGTAGTATGGGTTCCACTTCT 3102
DB 121 TTAATTTCTAAGAGGCTCTAAGAAACATGAATATCTCAGTAGTATGGGTTCCACTTCT 180
QY 3103 CCTCAATCCGTTGGATGAATAATTACTATGTGCCCTTAAGCACAATAATAGCTA 3157
DB 181 CCTCAATCCGTTGGATGAATAATTACTATGTGCCCTTAAGCACAATAATAGCTA 235
RESULT 13
ABN33794
ID ABN33794 standard; DNA; 60 BP.
XX ABN33794;
AC
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:6542.
DE
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200210449-A2.
PN
XX
XX 07-FEB-2002.
PD
```


XX 20-JUL-2001; 2001WO-IB001903.
PF 28-JUL-2000; 2000US-0221607P.
XX 02-MAY-2001; 2001US-0287724P.
XX (COMP-) COMPUGEN INC.
XX Shoshan A, Masserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
XX Example 1; SEQ ID NO 6542; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridizing selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at http://wipo.int/pub/published_pct_sequences
XX
XX Sequence 60 BP; 8 A; 20 C; 12 G; 20 T; 0 U; 0 Other;
SQ
Query Match 1.9%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.2e-18;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 571 TCTCTGAGCTTCGTCAAGTTTCATGCTGCTGCTCAAGATGCTACCTCTCTCACT 630
DB 1 TCTCTGAGCTTCGTCAAGTTTCATGCTGCTGCTGCTCAAGATGCTACCTCTCTCACT 60
RESULT 14
ABQ22410/c
ID ABQ22410 standard; DNA; 1026 BP.
XX
XX ABQ22410;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 9001.
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX WO200218632-A2.
XX

PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EP010074.
PF
XX 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
PR
XX (EPRI-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guettig D;
XX WPI; 2002-371829/40.
XX
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ41411 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
XX Sequence 1026 BP; 140 A; 128 C; 429 G; 329 T; 0 U; 0 Other;
SQ
Query Match 1.0%; Score 32; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1524 ACCCCCTGCGCCCAAGACTCGCGCGCTCGCA 1555
DB 915 ACCCCCTGCGCCCAAGACTCGCGCGCTCGCA 884
RESULT 15
ABQ22411
ID ABQ22411 standard; DNA; 1026 BP.
XX
XX ABQ22411;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 9002.
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX WO200218632-A2.
XX
XX 07-MAR-2002.
PD
XX 01-SEP-2001; 2001WO-EP010074.
PF

XX 01-SEP-2000; 2000DE-01043826.
 PR 05-SEP-2000; 2000DE-01044543.

XX (EPIC-) EPIGENOMICS AG.
 PA

PI Olek A, Piepenbrock C, Berlin K, Guetig D,
 DR WPI; 2002-371829/40.

XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's) and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX
 SQ Sequence 1026 BP; 329 A; 429 C; 128 G; 140 T; 0 U; 0 Other;

Query Match

1.0%; Score 32; DB 6; Length 1026;

Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTCGCCCAAGACTCGCGCGCTCGCA 1555

DB 112 ACCCCCTCGCCCAAGACTCGCGCGCTCGCA 143

Search completed: December 20, 2005, 17:22:43
 Job time : 1810 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 16:05:34 ; Search time 12030 Seconds
(without alignments)
12390.993 Million cell updates/sec

Title: US-10-649-400-3
Perfect score: 3186
Sequence: 1 cagcgtgattcctcaatlaa.....caccacaacaccttaagg 3186

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST:
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*
10: gb_esc10:*
11: gb_esc11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	777	24.4	1169	3	BQ067178
2	772	24.2	820	3	BI915063
3	754	23.7	957	5	BX441549
4	741	23.3	1044	3	BM922065
5	725	22.8	1116	5	BX418703
6	701	22.0	824	5	BX418656
7	694	21.8	768	8	DR001136
8	691	21.7	921	5	BX420367
9	625	19.6	945	3	BI914132
10	611	19.2	1028	3	BM548197
11	588	18.5	872	3	BI914330
12	581	18.2	583	3	BP311133
13	536	16.8	639	3	BI819019
14	531	16.7	740	3	BI917781
15	523	16.4	831	1	AM003023
16	508	15.9	534	5	BX114402
17	501	15.7	1154	3	BM807553
18	491	15.4	930	3	BI912693
19	454	14.2	582	3	BP228857
20	427	13.4	739	3	BI918597
21	426	13.4	779	3	BI820633
22	424	13.3	475	1	AM026293

C	23	419	13.2	606	1	AM590019	AM590019 hg29e09.x
	24	416	13.1	748	3	BI917426	BI917426 603184973
	25	415	13.0	772	3	BI823704	BI823704 603040821
	26	389	12.2	791	3	BI918360	BI918360 603182871
	27	385	12.1	583	3	BP229748	BP229748 BP229748
	28	380	11.9	463	1	AL045693	AL045693 DKF2P34B
C	29	359	11.3	541	1	AI201654	AI201654 q141f08.x
	30	350	11.0	468	10	CG464729	CG464729 KR1B.2D
	31	348	10.9	966	3	BI754056	BI754056 603027611
	32	343	10.8	425	10	AY419847	AY419847 Homo sapi
	33	335	10.5	386	1	AM160871	AM160871 au76g07.y
	34	334	10.5	910	7	CN643299	CN643299 ILUMIGEN
	35	334	10.5	910	7	CN805179	CN805179 ILUMIGEN
	36	328	10.3	395	8	H41824	H41824 Y007a05.r1
	37	322	10.1	507	8	R54288	R54288 YG78F01.r1
C	38	312	9.8	353	1	AM898907	AM898907 CM0-NN007
	39	312	9.8	363	1	AI815469	AI815469 au47f05.y
	40	289	9.1	991	3	BI757271	BI757271 603030777
	41	283	8.9	386	8	D56477	D56477 HMD430B05B
	42	281	8.8	463	8	H10108	H10108 YMD3d06.r1
	43	275	8.6	1015	3	BM921945	BM921945 AGENCOURT
	44	260	8.2	389	8	M62287	M62287 EST00360.Hi
C	45	255	8.0	385	2	BF059564	BF059564 7k62g05.x

ALIGNMENTS

RESULT 1
LOCUS BQ067178 1169 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6767481 NIH_MGC_115 Homo sapiens CDNA IMAGE:5751547
5', mRNA sequence.
ACCESSION BQ067178
VERSION BQ067178.1 GI:19896224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
CDNA Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12784 row: e column: 20
High quality sequence start: 16
High quality sequence stop: 685.
Location/Qualifiers
1..1169
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5751547"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector: pcmv-SPORE6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age 69; library is male lung, age 27; and 1 male testis, age 69; library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. library is normalized and enriched for full-length clones and was constructed by C.

FEATURES

source

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

Query Match 24.4%; Score 777; DB 3; Length 1169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 777; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

826 CCTACCACAGACGATACGGCCCTCTTACCAAGAGCTCCCTTCTACCAAGTTCTCTCCA 885
36 CCTACCACAGACGATACGGCCCTCTTACCAAGAGCTCCCTTCTACCAAGTTCTCTCCA 95
886 CCCAGCCGGGGCTGTGTGCCCCGCAAGACAGATGATCTGTGCAACAGAGCCCTTTGGC 945
96 CCCAGCCGGGGCTGTGTGCCCCGCAAGACAGATGATCTGTGCAACAGAGCCCTTTGGC 155
946 TGAATTTTACCGGACCAACAGAGATGATCATCAACCAAGGGAAGGCGCATGTTTC 1005
156 TGAATTTTACCGGACCAACAGAGATGATCATCAACCAAGGGAAGGCGCATGTTTC 215
QY 1006 CTTTCTTAAAGTTTAACTTTCTGTCTGTGATCCCAAGGCTCATTAACAATATTTTGTGG 1065
Db 216 CTTTCTTAAAGTTTAACTTTCTGTCTGTGATCCCAAGGCTCATTAACAATATTTTGTGG 275
QY 1066 ATGTGATTTTGGCGGATCCCAATCACTGAGAGGTTTCAAGAGCAAAATGGTTCCTTGG 1125
Db 276 ATGTGATTTTGGCGGATCCCAATCACTGAGAGGTTTCAAGAGCAAAATGGTTCCTTGG 335
QY 1126 GCAAAACCGGACCAATGTGCAAGAAATCGGGTCTATATGATCCGATTCGCCCAACA 1185
Db 336 GCAAAACCGGACCAATGTGCAAGAAATCGGGTCTATATGATCCGATTCGCCCAACA 395
QY 1186 CTGGGGGCTCACTGTGATCGCCCAAGAAATCTCTTTGAAAAATTAAACTTACGACAACA 1245
Db 396 CTGGGGGCTCACTGTGATCGCCCAAGAAATCTCTTTGAAAAATTAAACTTACGACAACA 455
QY 1246 AAGAGCTTCAATATAAATGAGAGATGGTGTCTTACAGTCTTGGCAAGTACAGC 1305
Db 456 AAGAGCTTCAATATAAATGAGAGATGGTGTCTTACAGTCTTGGCAAGTACAGC 515
QY 1306 CCGGCTCGATGTGTGAAAGTGAACGAGAGCGGACGAGGACACTAGCCAGCCGGCC 1365
Db 516 CCGGCTCGATGTGTGAAAGTGAACGAGAGCGGACGAGGACACTAGCCAGCCGGCC 575
QY 1366 GCGTGCAGACCTTCACTTCCCTGAGACTCACTTATCATCGCCCTCAACCGCTTCAAGACA 1425
Db 576 GCGTGCAGACCTTCACTTCCCTGAGACTCACTTATCATCGCCCTCAACCGCTTCAAGACA 635
QY 1426 CGGATATTACAACTGAATAGATCAACAACCTTTTGAAGAAGATTTCGGATTAAT 1485
Db 636 CGGATATTACAACTGAATAGATCAACAACCTTTTGAAGAAGATTTCGGATTAAT 695
QY 1486 ATGACGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTTGGCCCAACGACTGGC 1545
Db 696 ATGACGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTTGGCCCAACGACTGGC 755
QY 1546 CGCGCTCGAGATCGTGTGCCCCGGGGCCCGCTACGCAATGGCCGCTTTCTCTGACAG 1602
Db 756 CGCGCTCGAGATCGTGTGCCCCGGGGCCCGCTACGCAATGGCCGCTTTCTCTGACAG 812

RESULT 2
BI915063 820 bp mRNA linear EST 16-OCT-2001
LOCUS BI915063 603177258F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241542 5',
DEFINITION mRNA sequence.
ACCESSION BI915063
VERSION BI915063.1 GI:16199126
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM11609 row: c column: 15
High quality sequence stop: 816.

FEATURES
source
Location/Qualifiers
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5241542"
/lab_host="DH10B"
/clone_id="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dt primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 24.2%; Score 772; DB 3; Length 820;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 GAGCGGAGGGCGAGTGTTCAGGTTCTAGAGCTATGAGCTGGAGCACTGCTTCTCTCT 325
Db 49 GAGCGGAGGGCGAGTGTTCAGGTTCTAGAGCTATGAGCTGGAGCACTGCTTCTCTCT 108
QY 330 TCTATCATGCTCTCCAGAAATTTCTCAATGTGAGGAGCACTACCAATTGAGGGGA 389
Db 109 TCTATCATGCTCTCCAGAAATTTCTCAATGTGAGGAGCACTACCAATTGAGGGGA 168
QY 390 TCCGAGCTTGTCTTGAAGATCATCCATTATCTGACCACTGACCAACTGGAGAGAAT 449
Db 169 TCCGAGCTTGTCTTGAAGATCATCCATTATCTGACCACTGACCAACTGGAGAGAAT 228
QY 450 TCACCTTTGAAAAAATTACAGGGGATGACGAATAGTCAATACAGACAATTTTCT 509
Db 229 TCACCTTTGAAAAAATTACAGGGGATGACGAATAGTCAATACAGACAATTTTCT 288
QY 510 GACTCCAAAGACTACACAGGGGAGCTCCAGAGAAATTAACCTCTCTGCTTGGAGGG 569
Db 289 GACTCCAAAGACTACACAGGGGAGCTCCAGAGAAATTAACCTCTCTGCTTGGAGGG 348
QY 570 GTCTCTGAGCTTCTGTCAGATTTGATGAGCTTGTGCTGCTGCAATGCTCTCTCTAG 629
Db 349 GTCTCTGAGCTTCTGTCAGATTTGATGAGCTTGTGCTGCTGCAATGCTCTCTCTAG 408
QY 630 TCCAGCCAGCCACAGTCTGCGGCACTGCTCCAGTGCATGTTCCGCTTACCCGGCCAG 689
Db 409 TCCAGCCAGCCACAGTCTGCGGCACTGCTCCAGTGCATGTTCCGCTTACCCGGCCAG 468
QY 690 CAGGACCGGGGACCCCGCTTCTCATGCGGACCCCTGAGCTACATGGCCACAC 749
Db 469 CAGGACCGGGGACCCCGCTTCTCATGCGGACCCCTGAGCTACATGGCCACAC 528
QY 750 CCGGTATCAACCAAGAGCTTACAGAGCTCTGTCCAATCTCTGCGGACAGGATAC 809

```

Db      529  |||||
CGGTCATCACCAAGAGCTTACCAAGCTCTCTGTCTCACTCTCTGCGCCAGGATAC 588
Oy      810  |||||
CCACGCGCGGCTACCCCTTACCCAGAGTACGGCACTCTTACCAAGAGCTCTGTC 869
Db      589  |||||
CCCAAGCGCGGCTTACCCCTTACCCAGAGTACGGCACTCTTACCAAGAGCTCTGTC 648
Oy      870  |||||
TACCAATTCCTCTCAACCAAGCGGCGTGTGCTCCGCGCAAGCAAGTGTACTGTGC 929
Db      649  |||||
TACCAATTCCTCTCAACCAAGCGGCGTGTGCTCCGCGCAAGCAAGTGTACTGTGC 708
Oy      930  |||||
AACAGAGCCCTTGTGCTGAATTTACCGGCAACAAAGAGATGATCATCAACAAAG 989
Db      709  |||||
AACAGAGCCCTTGTGCTGAATTTACCGGCAACAAAGAGATGATCATCAACAAAG 768
Oy      990  |||||
GGAAGGCGCATGTTCTTTTAAATTTCATTTCTGTCTGATCCCA 1041
Db      769  |||||
GGAAGGCGCATGTTCTTTTAAATTTCATTTCTGTCTGATCCCA 820

RESULT 3
BX441549 957 bp mRNA linear EST 05-MAY-2004
LOCUS BX441549 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF018YJ12 5-PRIME, mRNA sequence.
ACCESSION BX441549
VERSION BX441549.2 GI:47038398
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 957)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31018481.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DF018DE060P1&c=1851.f.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF018YJ12"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1lb="Homo sapiens FETAL BRAIN"
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 23.7%; Score 754; DB 5; Length 957;
Best local Similarity 99.9%; Pred. No. 0;
Matches 874; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy      387  |||||
GGATCCGAGCTTGTGCGAGTATCATTCATTCGACCACTGACCACTCTGGAGAGA 446
Db      8  |||||
GGATCCGAGCTTGTGCGAGTATCATTCATTCGACCACTGACCACTCTGGAGAGA 67

```

```

Oy      447  |||||
AGTTACCTTTGAAAAAATTACAGGGGATGACGATAGTCAGATACAGCAATTTT 506
Db      68  |||||
AGTTACCTTTGAAAAAATTACAGGGGATGACGATAGTCAGATACAGCAATTTT 127
Oy      507  |||||
CTGACTTCAAGGATCAACAGGAGGATCTCAGAGAAATTAATCTCTCTCTTGGAC 566
Db      128  |||||
CTGACTTCAAGGATCAACAGGAGGATCTCAGAGAAATTAATCTCTCTCTTGGAC 187
Oy      567  |||||
GGGGTCTGAGCTTCTGTCACAGTTTGATGGCTCTGCTGAGATGCTCTCTCTCT 626
Db      188  |||||
GGGGTCTGAGCTTCTGTCACAGTTTGATGGCTCTGCTGAGATGCTCTCTCTCT 247
Oy      627  |||||
CAGTCCAGCGGCAAGTCTGGGCGCACTGCTCCAGTGGCATGTTCCGTACCCGCG 686
Db      248  |||||
CAGTCCAGCGGCAAGTCTGGGCGCACTGCTCCAGTGGCATGTTCCGTACCCGCG 307
Oy      687  |||||
CAGACGAGACCGGCGCAACCCGCTTCTCATCGGAGCCCTAGCCGCTACATGCGC 746
Db      308  |||||
CAGACGAGACCGGCGCAACCCGCTTCTCATCGGAGCCCTAGCCGCTACATGCGC 367
Oy      747  |||||
CACCAGTCAATCAACAGGAGCTTCAAGAGCTTCTCTGCTGCTGCTGCTGCTGCT 806
Db      368  |||||
CACCAGTCAATCAACAGGAGCTTCAAGAGCTTCTCTGCTGCTGCTGCTGCTGCT 426
Oy      807  |||||
TACCCAGAGCGGCTACCCCTTACCCAGAGATGAGGCACTCTTACCAAGAGCTCG 866
Db      427  |||||
TACCCAGAGCGGCTACCCCTTACCCAGAGATGAGGCACTCTTACCAAGAGCTCG 486
Oy      867  |||||
TTTTCACAGTTCTCTCCACCAAGCGGCGCTGTGCTGCGGCAAGCAAGGTACTCG 926
Db      487  |||||
TTTTCACAGTTCTCTCCACCAAGCGGCGCTGTGCTGCGGCAAGCAAGGTACTCG 546
Oy      927  |||||
TGCAACAGGCGCTTGTGCTGAATTTTCAACGCGGCAACAAAGAGATATATACCAA 986
Db      547  |||||
TGCAACAGGCGCTTGTGCTGAATTTTCAACGCGGCAACAAAGAGATATATACCAA 606
Oy      987  |||||
CAGGAGAGCGGCTTCTCTTTTAAATTTCATTTCTGCTGATCCAGGCT 1046
Db      607  |||||
CAGGAGAGCGGCTTCTCTTTTAAATTTCATTTCTGCTGATCCAGGCT 666
Oy      1047  |||||
CATTCATATTTTGTGATGATGATTTTGGCGATGCCAATCATCTGAGGTTTCAAGA 1106
Db      667  |||||
CATTCATATTTTGTGATGATGATTTTGGCGATGCCAATCATCTGAGGTTTCAAGA 726
Oy      1107  |||||
GGCAATGAGTCTCTTGGCGGCAAGCGGCAACCAATGTGCAAGAAATCGGCTATATG 1166
Db      727  |||||
GGCAATGAGTCTCTTGGCGGCAAGCGGCAACCAATGTGCAAGAAATCGGCTATATG 786
Oy      1167  |||||
CATCGGATTCGCCCAACATGCGGCTCATGAGTGGCCCAAGAAATCTTTTGGAAA 1226
Db      787  |||||
CATCGGATTCGCCCAACATGCGGCTCATGAGTGGCCCAAGAAATCTTTTGGAAA 846
Oy      1227  |||||
TTAAACTTACGACCAAGAGAGCTTCAATAA 1261
Db      847  |||||
TTAAACTTACGACCAAGAGAGCTTCAATAA 881

RESULT 4
BM922065 1044 bp mRNA linear EST 12-MAR-2002
LOCUS BM922065 AGENCOURT 6707015 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754122
DEFINITION 5', mRNA sequence.
ACCESSION BM922065
VERSION BM922065
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1044)
NIH-MGC http://mgs.nci.nih.gov/.

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM12791 row: a column: 03
High quality sequence stop: 694.
Location/Qualifiers

FEATURES
source

```
1..1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754122"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      23.3%; Score 741; DB 3; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 202 AGTGGCTTCTGTCTAGTGAAGGGGCTGTGATTTCTAGTTATGTAATAGACTTTA 261
Db 23 AGGCTTCTGTCTAGTGAAGGGGCTGTGATTTCTAGTTATGTAATAGACTTTA 82
Oy 262 AAAACGAGGAGCGGAGGCGAGTGTTCAGTTCAGACTTAGAGCTGAGGAGCACTGCC 321
Db 83 AAAACGAGGAGCGGAGGCGAGTGTTCAGTTCAGACTTAGAGCTGAGGAGCACTGCC 142
Oy 322 TTTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGCACTCACTT 381
Db 143 TTTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGCACTCACTT 202
Oy 382 CAGGGGGAGTCGAGCTGTCTGACAGATCATCCCATTTCTCGACCACTGACAACTGG 441
Db 203 CAGGGGGAGTCGAGCTGTCTGACAGATCATCCATTTCTCGACCACTGACAACTGG 262
Oy 442 AGAGAAGTTCACCTTTGAAAAAATTACAGAGGGGAGTGAAGATCAGTCAGATAGACACA 501
Db 263 AGAGAAGTTCACCTTTGAAAAAATTACAGAGGGGAGTGAAGATCAGTCAGATAGACACA 322
Oy 502 ATTTTCTGACTCCAAAGACTCACAGGGGAGCTCCAGAAAGTAACTCTCTCCTGCT 561
Db 323 ATTTTCTGACTCCAAAGACTCACAGGGGAGCTCCAGAAAGTAACTCTCTCCTGCT 382
Oy 562 TGGAGCGGGGCTCTGAGCTTCGTCAGTTCGATGGGCTCTGTCAGATGGCTACCTCC 621
Db 383 TGGAGCGGGGCTCTGAGCTTCGTCAGATGGGCTCTGTCAGATGGCTACCTCC 442
Oy 622 TCTCTCAGTCAGCCAGCCAGCAGTCTGCGGCACTGCTCCAGTGCATGTTCCGCTAC 681
Db 443 TCTCTCAGTCAGCCAGCCAGCAGTCTGCGGCACTGCTCCAGTGCATGTTCCGCTAC 502
Oy 682 CCGGCGAGCACGGACCGGCGCACCCTTCTTCATCGGCAAGCCTTACCCGCTATACG 741
Db 503 CCGGCGAGCACGGACCGGCGCACCCTTCTTCATCGGCAAGCCTTACCCGCTATACG 562
```

```
Oy 742 CCACACACCGGTCATCAACCAAGGAGCTTACACAGCTCTCTGTCCACTCTCTCGCGC 801
Db 563 CCACACACCGGTCATCAACCAAGGAGCTTACACAGCTCTCTGTCCACTCTCTCGCGC 622
Oy 802 AGGATACCCCAACCGGCGGCTACCTCTTACCAACAGCATGAGGCACTCTTACCAAGAG 861
Db 623 AGGATACCCCAACCGGCGGCTACCTCTTACCAACAGCATGAGGCACTCTTACCAAGAG 682
Oy 862 CTCGGTCTTACCACTTCCCTCCACCAAGCGGGGCTGTGCCCGGCAAGACAGAGTGT 921
Db 683 CTCGGTCTTACCACTTCTCTCCACCAAGCGGGGCTGTGCCCGGCAAGACAGAGTGT 742
Oy 922 ACCTGTGCAACAGGCGCCCTTT 942
Db 743 ACCTGTGCAACAGGCGCCCTTT 763
```

RESULT 5
BX418703
LOCUS
DEFINITION BX418703 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION BX418703
VERSION BX418703.2 GI:46957805
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1116)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30769503.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?cs=SDUF008D10P1&c=1851.f>.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENTFEATURES
source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF0081F20"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the PCMVSPORT 6
vector. Library was not normalized."
```

ORIGIN

```
Query Match      22.8%; Score 725; DB 5; Length 1116;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 273 CGGAGGGCGAGTGTTCAGTTCAGAGCTTACAGAGCTGAGCACTGCTTCTCTCTCT 332
Db 9 CGGAGGGCGAGTGTTCAGTTCAGAGCTTACAGAGCTGAGCACTGCTTCTCTCTCT 68
Oy 333 ATCATGCTCTTCAAGAAATTTCTCAATGTGAGCAGCACTCAATTCAGGCGGATCC 392
```


DB	LOCUS	DEFINITION	ACCESSION VERSION KEYWORDS	SOURCE ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL COMMENT
Db	544	CCACACCCGGTCATTCACCAACGAGCCTTCTGTCAACTCCTGCCTCA					
Oy	803	GGGATATCCCAACGCGCGGCTACCCCTTACCACAGCAGTACGGCCTCTACCAAGAGC					
Db	604	-GGATATCCCAACGCGCGGCTACCCCTTACCACAGCAGTACGGCCTCTACCAAGAGC					
Oy	863	TCGGTTTACACAGTTCTCTCTCCACCCAGCGGGGCTGGTGCCTGGCAACAGCAGTGTGA					
Db	663	TCGGTTTACACAGTTCTCTCTCCACCCAGCGGGGCTGGTGCCTGGCAACAGCAGTGTGA					
Oy	923	CCTGTGCAACAGCGCCCTTTGGCTGAATTTTACACCGGCAACAAAGGAGATGATCATAC					
Db	723	CCTGTGCAACAGCGCCCTTTGGCTGAATTTTACACCGGCAACAAAGGAGATGATCATAC					
Oy	983	CAACAGGGAAGCGCATGTTTCTTTTAACTTTTAACT 1024					
Db	783	CAACAGGGAAGCGCATGTTTCTTTTAACTTTTAACT 824					
RESULT 7	DR001136	768 bp mRNA linear EST 17-MAY-2005					
LOCUS	TC115971	Human fetal brain, large insert, PCMV expression library					
DEFINITION	DR001136	DR001136.1 GI:66261009					
ACCESSION	DR001136	EST.					
VERSION	DR001136.1	GI:66261009					
KEYWORDS		EST.					
SOURCE	DR001136	768 bp mRNA linear EST 17-MAY-2005					
ORGANISM	DR001136	768 bp mRNA linear EST 17-MAY-2005					
REFERENCE	DR001136	768 bp mRNA linear EST 17-MAY-2005					
AUTHORS	DR001136	768 bp mRNA linear EST 17-MAY-2005					
TITLE	DR001136	768 bp mRNA linear EST 17-MAY-2005					
JOURNAL	DR001136	768 bp mRNA linear EST 17-MAY-2005					
COMMENT	DR001136	768 bp mRNA linear EST 17-MAY-2005					

ORIGIN	verification of full-length genes"
Query Match	21.8%; Score 694; DB 8; Length 768;
Best Local Similarity	100.0%; Pred. No. 0;
Matches	694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	203 GTGCTTCTGTCTAGTGAAGGGGCTGTGTGATTTCTAGTTTATGATTAATAGACTTTAA 262
DB	9 GTGCTTCTGTCTAGTGAAGGGGCTGTGTGATTTCTAGTTTATGATTAATAGACTTTAA 68
QY	263 AAACGAGGGAAGGAGGGGCGAGTGTTCAGAGTTCTAGAGCTATGAGCTGTGAGCACTGCT 322
DB	69 AAACGAGGGAAGGAGGGGCGAGTGTTCAGAGTTCTAGAGCTATGAGCTGTGAGCACTGCT 128
QY	323 TTCTCTTCTATCATGCTCTTCAAGAAATTTCTCAATGTGAGCAGCAGCTAACCAATTC 382
DB	129 TTCTCTTCTATCATGCTCTTCAAGAAATTTCTCAATGTGAGCAGCAGCTAACCAATTC 188
QY	383 AGCGGATCGAGCTTGTCTTGACAGATCATCCATTATCTCGACCACTGACCAACTGGA 442
DB	189 AGCGGATCGAGCTTGTCTTGACAGATCATCCATTATCTCGACCACTGACCAACTGGA 248
QY	443 GAGAAGTTCACCTTTGAAAAAATTCACAGGGGGATAGCAATCAGTCAGATTCAGACAA 502
DB	249 GAGAAGTTCACCTTTGAAAAAATTCACAGGGGGATAGCAATCAGTCAGATTCAGACAA 308
QY	503 TTTTCTGACTCCAAAGACTCACCAGGGGAGCTTCAGAGAAATMACTCTCTCTGTCTT 562
DB	309 TTTTCTGACTCCAAAGACTCACCAGGGGAGCTTCAGAGAAATMACTCTCTCTGTCTT 368
QY	563 GAGCGGGGTCTCTGAGCTTGTCTCAAGTTTGATGCTCTGTCTGACAGATGGCTACCTCT 622
DB	369 GAGCGGGGTCTCTGAGCTTGTCTCAAGTTTGATGCTCTGTCTGACAGATGGCTACCTCT 428
QY	623 CTCTCAGTCCAGCCAGCCACAGCTGTGGGGCACGTCTCCAGTGCATGTTCCTGATACC 682
DB	429 CTCTCAGTCCAGCCAGCCACAGCTGTGGGGCACGTCTCCAGTGCATGTTCCTGATACC 488
QY	683 CGGCGAGCAGGAGCCGGGCGACCCCGGCTTCTCATGAGGAGCCCTAGCCGCTACATGAGC 742
DB	489 CGGCGAGCAGGAGCCGGGCGACCCCGGCTTCTCATGAGGAGCCCTAGCCGCTACATGAGC 548
QY	743 CCACCACCCGGTATCAACAAAGGAGCTTCAACAGCTTCTGTCAACTCTCTGCGCGCA 802
DB	549 CCACCACCCGGTATCAACAAAGGAGCTTCAACAGCTTCTGTCTGCGCGCGCA 608
QY	803 GGGATATCCCAAGGCGGCTACCCCTTACCAACAGCTAAGGAGGAGGAGGAGGAGGAGG 862
DB	609 GGGATATCCCAAGGCGGCTACCCCTTACCAACAGCTAAGGAGGAGGAGGAGGAGGAGG 668
QY	863 TCCGTTCTACCAAGTTCCTTCCACCCAGCCGCGGG 896
DB	669 TCCGTTCTACCAAGTTCCTTCCACCCAGCCGCGGG 702
RESULT 8	
LOCUS	BX420367 921 bp mRNA linear EST 01-MAY-2004
DEFINITION	BX420367 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION	CS0DF0277H06 5-PRIME, mRNA sequence.
VERSION	BX420367
KEYWORDS	BX420367.2 GI:46932405
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
JOURNAL	Homidae; Homo.
	1 (bases 1 to 921)
	Lj. W. B., Gruber C., Jesse, J. and Polayes, D.
	Full-length cDNA libraries and normalization
	Unpublished (2001)

COMMENT

On May 13, 2003 this sequence version replaced gi:30650954.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 1851.f
 For more information about this cluster see
 http://www.genoscope.cns.fr/cdna?c=CS0DF027DD03QPI&c=1851.f.

FEATURES

source

1..921
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF027YH06"
 /tissue="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (dt) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN

Query Match 21.7%; Score 691; DB 5; Length 921;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 861; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 358 ATGTGAGGAGGAGTACCAATTCAGGAGGAGTCCAGTCTTCTGACAGATACCA 417
DB 1 ATGTGAGGAGGAGTACCAATTCAGGAGGAGTCCAGTCTTCTGACAGATACCA 60
QY 418 TTATCTGACACCTGACAACTGAGAGAGTTCACCTTTGAAAAAATTTACAGGGGGA 477
DB 61 TTATCTGACACCTGACAACTGAGAGAGTTCACCTTTGAAAAAATTTACAGGGGGA 120
QY 478 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
DB 121 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 538 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597
DB 181 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 598 GCTCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
DB 241 GCTCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 658 CTCCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 717
DB 301 CTCCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 718 TCGGAGACCTTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
DB 361 TCGGAGACCTTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 778 GCCTCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 837
DB 421 GCCTCTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479
QY 838 AGTACGGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 897
DB 480 AGTACGGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
QY 898 TGTGTCCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
DB 540 TGTGTCCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599

```

```

QY 958 GGACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
DB 600 GGACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
QY 1018 TTACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
DB 660 TTACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
QY 1078 CGGATCCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
DB 720 CGGATCCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
QY 1138 CCAATGTCAGAGAAATCGGCTATATGATCGGATTCGATCCCACTGGGCTCACT 1197
DB 780 CCAATGTCAGAGAAATCGGCTATATGATCGGATTCGATCCCACTGGGCTCACT 839
QY 1198 GGATGCGCCAGAAATCTCTTTT 1220
DB 840 GGATGCGCCAGAAATCTCTTTT 862

```

RESULT 9
 B1914132 945 bp mRNA linear EST 16-OCT-2001
 LOCUS 603182077F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246298 5',
 DEFINITION mRNA sequence.
 ACCESSION B1914132 GI:16178444
 VERSION B1914132.1 GI:16178444
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 945)
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LRAM11621 row: 1 column: 19
 High quality sequence stop: 711.
 Location/Qualifiers

FEATURES

source

1..945
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246298"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb. Insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 19.6%; Score 625; DB 3; Length 945;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 675; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 294 TCTAGAGCTATGACGTGGAGCACTGCTCTCTCTCTATCATCTCTCCAGAAATTT 353
DB 14 TTTAAGCTATGACGTGGAGCACTGCTCTCTCTCTCTATCATCTCTCCAGAAATTT 73
OY 354 CTCAAATGAGAGCAAGCTACCAATTCAGGCGGATCCGAGCTTGTCTTGACGATCAT 413
DB 74 CTCAATGTGAGAGCAAGCTACCAATTCAGGCGGATCCGAGCTTGTCTTGACGATCAT 133
OY 414 CCCATTATTCGACCACTGACAACTTGAGAGAGAAATTCCTTTGAAAAAATTCACGAG 473
DB 134 CCCATTATTCGACCACTGACAACTTGAGAGAGAAATTCCTTTGAAAAAATTCACGAG 193
OY 474 GGGATGACGAATCATGTCAGATACAGCAATTTTCTGACTCCAGAGACTCACAGGGAGC 533
DB 194 GGGATGACGAATCATGTCAGATACAGCAATTTTCTGACTCCAGAGACTCACAGGGAGC 253
OY 534 GTCCAGAGAGTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 593
DB 254 GTCCAGAGAGTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 313
OY 594 GATGCTCTGCTGCAAGATGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 653
DB 314 GATGCTCTGCTGCAAGATGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 373
OY 654 ACTGCTCCAGTGCATGTTCCCTGATCCCGGCGACAGACGAGCGGACCCCGCTTC 713
DB 374 ACTGCTCCAGTGCATGTTCCCTGATCCCGGCGACAGACGAGCGGACCCCGCTTC 433
OY 714 TCCATGCGAGCCCTGACCGCTACATGAGCCACACCCGCTCATCAACAGAGACTTAC 773
DB 434 TCCATGCGAGCCCTGACCGCTACATGAGCCACACCCGCTCATCAACAGAGACTTAC 493
OY 774 AACAGCTCTGCTGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 833
DB 494 AACAGCTCTGCTGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 553
OY 834 CAGCATGACGACACTCTCTACCAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 893
DB 554 CAGCATGACGACACTCTCTACCAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 613
OY 894 GGGCTGCTGCTGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 953
DB 614 GGGCTGCTGCTGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 673
OY 954 CACCGGACCAAAACG 969
DB 674 CACCGGACCAAAACG 689

RESULT 10
BMS48197 1028 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6571595 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5733882
DEFINITION 5', mRNA sequence.
ACCESSION BMS48197
VERSION BMS48197.1 GI:18782590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1028)
NIH-MGC http://mgi.mc.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

```

FEATURES
Source
1..1028
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5733882"
/tissue_type="hippocampus"
/lab_host="DH108"
/clone_1id="NIH_MGC_124"
/notes="Organ: brain; Vector: pCMV-Sport6; Site 1: EcorV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dt primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

ORIGIN
Query Match 19.2%; Score 611; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred. No. 1.3e-310;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 451 CACCTTTGAAAAAATTAACAGGGGGATGACGATTCGATACGACAAATTTCTCTG 510
DB 79 CACCTTTGAAAAAATTAACAGGGGGATGACGATTCGATACGACAAATTTCTCTG 138
OY 511 ACTCCAAGAGACTCACAGAGGGAGCGTCCAGAGAAAGTAATCTCTCTCTCTCTCTCTCTG 570
DB 139 ACTCCAAGAGACTCACAGAGGGAGCGTCCAGAGAAAGTAATCTCTCTCTCTCTCTCTCTG 198
OY 571 TCTCTGAGCTTGTGACAGTTTGGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 630
DB 199 TCTCTGAGCTTGTGACAGTTTGGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 258
OY 631 CCAAGCCAGCCACATCTGCGGCGCACTGCTCCAGATGCAATGTTCCGCTACCCCGGCGAGC 690
DB 259 CCAAGCCAGCCACATCTGCGGCGCACTGCTCCAGATGCAATGTTCCGCTACCCCGGCGAGC 318
OY 691 ACGGACCGGCGACCCCGCTTCTCTCATCGGACGCTTACCGGCTTACGCTTACGCTTACG 750
DB 319 ACGGACCGGCGACCCCGCTTCTCTCATCGGACGCTTACCGGCTTACGCTTACGCTTACG 378
OY 751 CGGTCAATCAACCAAGAGAGCTTACCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 810
DB 379 CGGTCAATCAACCAAGAGAGCTTACCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 438
OY 811 CCAAGCGCGGCTTACCCCTTACCAAGAGATGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
DB 439 CCAAGCGCGGCTTACCCCTTACCAAGAGATGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 498
OY 871 ACAAGTTCTCTTCCACCAAGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 930
DB 499 ACAAGTTCTCTTCCACCAAGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 558
OY 931 GAAGCGGAGATGTTCTTTTAAAGTTTAAATTTCTGCTCTGATCCCAAGGCTGATT 1050
DB 619 GAAGCGGAGATGTTCTTTTAAAGTTTAAATTTCTGCTCTGATCCCAAGGCTGATT 678
OY 1051 ACAATATTTT 1061
DB 679 ACAATATTTT 689

RESULT 11

```

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12738 row: e column: 19
High quality sequence stop: 623.
Location/Qualifiers

B1914330 872 bp mRNA linear EST 17-OCT-2001
 LOCUS 603182311P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246210.5
 DEFINITION mRNA sequence.
 ACCESSION B1914330
 KEYWORDS EST.
 VERSION B1914330.1 GI:16198838
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 872)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgsabrs-rc@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM1621 row: f column: 03
 High quality sequence stop: 853.
 Location/Qualifiers
 1..872
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246210"
 /lab_host="DH10B"
 /lab_host="NIH_MGC_121"
 /notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NciI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."
 ORIGIN
 Query Match 18.5%; Score 588; DB 3; Length 872;
 Best Local Similarity 100.0%; Pred. No. 1.8e-298;
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 405 CAACAGCCTCTCTTCCAACTCTCGCCGACGAGATACCCACGCGCGCTACCCCTACCC 464
 Oy 833 ACAGCAGTACGGCCCACTCTTACCAAGAGCTCCGTTCTACAGTTCCTCCACCAAGCC 892
 Db 465 ACAGCAGTACGGCCCACTCTTACCAAGAGCTCCGTTCTACAGTTCCTCCACCAAGCC 524
 Oy 893 GGGGTGTGTGCGCCGCAAGACAGTGTACTGTGCAACAGGCCCTTTGGCTGAATTT 952
 Db 525 GGGGTGTGTGCGCCGCAAGACAGTGTACTGTGCAACAGGCCCTTTGGCTGAATTT 584
 Oy 953 TCACCGGACCAACCGAGATGATCATCAACCAACAGGAAAGCGCATGTTTCTTTT 1012
 Db 585 TCACCGGACCAACCGAGATGATCATCAACCAACAGGAAAGCGCATGTTTCTTTT 644
 Oy 1013 AAGTTTAACTTTCTGTCTCGATCCACGCGTCATTACAAATTTT 1060
 Db 645 AAGTTTAACTTTCTGTCTCGATCCACGCGTCATTACAAATTTT 692
 RESULT 12
 Bp311133
 LOCUS Bp311133 Sugano cDNA library, brain Homo sapiens cDNA clone
 DEFINITION NR04714, mRNA sequence.
 ACCESSION Bp311133
 VERSION Bp311133.1 GI:52240108
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 583)
 Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 15342556
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NR04714"
 /tissue_type="brain"
 /clone_lib="Sugano cDNA library, brain"
 ORIGIN
 Query Match 18.2%; Score 581; DB 3; Length 583;
 Best Local Similarity 100.0%; Pred. No. 9.1e-295;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	430	TTGACAACTGGAGGAAGTCACTTTGAAAAAAATTACAGGGGGATGACGATTCGT	489
Db	243	CTGACAACTGGAGGAAGTTACCTTTGAAAAAAATTACAGGGGGATGACGATTCGT	302
Qy	490	CAGATACAGACAATTTTCTCTGACCTTCCAAAGAACTCACAGGGGACGTCCAGAGAATTAAC	549
Db	303	CAGATACAGACAATTTTCTCTGACTCCAAAGAACTCACAGGGGACGTCCAGAGAATTAAC	362
Qy	550	TCTCTCCGTCTTTGGACGGGGCTCTGAGCTCGACAGTTCGATGGCTGTGCTGAG	609
Db	363	TCTCTCCTGTCTTTGGACGGGGTCTCTGAGCTTCGACAGTTTCGATGGCTGTGCTGAG	422
Qy	610	ATCGCTACCTCTCTCTCAGTTCAGTCAGCCAGCCACAGTCTTGGGACACTGCTCCAGTCCCA	669
Db	423	ATCGCTACCTCTCTCTCAGTTCAGTCAGCCAGCCACAGTCTGGGGCACTGCTCCCAAGTCCCA	482
Qy	670	TGTTCCCGTACCCCGGAGCCAGACGGAGCACCAGCCCTTTCGATCGGACAGCCCTA	729
Db	483	TGTTCCCGTACCCCGGAGCCAGACCGAGCCACCAGCCCTTTCGATCGGACAGCCCTA	542
Qy	730	GCGGCTACATGGCCCAACCACTCGGTGATATCAACCAAGGAGCC	770
Db	543	GCGGCTACATGGCCCAACCACTCGGTATATCAACCAAGGAGCC	583

RESULT 13					
BI819019					
LOCUS					
DEFINITION	BI819019	639 bp	mRNA	linear	EST 04-OCT-2001
	60303341.F1				
	Homo sapiens CDNA clone IMAGE:5174290 5', mRNA sequence.				

ACCESSION	BI19019
VERSION	BI19019.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 639)
AUTHORS	NIH-MGC http://mgc.nhl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cs9abbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLMW1434 row: a column: 11
High quality sequence stop: 639.

FEATURES	Location/Qualifiers
source	1. .639

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174290"
/lab_host="DH10B"
/clone_1fb="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

```

ORIGIN

Query Match 16.8%; Score 536; DB 3; Length 639;
Best Local Similarity 100.0%; Pred. No. 5.6e-271;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	473	GGGATATACGAATAGTCAGATACAGACAATTTTTCGATCTCCAAAGACATCCAGGGGA	532
Db	104	GGGATATACGAATAGTCAGATACAGACAATTTTTCGATCTCCAAAGACATCCAGGGGA	163
QY	533	CGTCCAGAGAAAGTAACTCTCTCTGTCTTTGGACGGGGTCTCTGAGCTTGTCACAGTTT	592
Db	164	CGTCCAGAGAAAGTAACTCTCTCTGTCTTTGGACGGGGTCTCTGAGCTTGTCACAGTTT	223
QY	593	CGATGGCTCTGCTGCATGCGATGCGTACCTCTCTCACTTCAGCGAGCCACAGTCTGGGCG	652
Db	224	CGATGGCTCTGCTGCATGCGATGCGTACCTCTCTCACTTCAGCGAGCCACAGTCTGGGCG	283
QY	653	CACGTCTCCAGTGCATGTTTCCCGTACCCCGGACAGCACGGACCGGCGACCCCGCTT	712
Db	284	CACGTCTCCAGTGCATGTTTCCCGTACCCCGGACAGCACGGACCGGCGACCCCGCTT	343
QY	713	CTCCATCGGCGAGCCCTTAGCCGCTACATAGGCCACACCCGGTATATACAAACGGAAGCTA	772
Db	344	CTCCATCGGCGAGCCCTTAGCCGCTACATAGGCCACACCCGGTATATACAAACGGAAGCTA	403
QY	773	CAACAGCCTCTGTCTCAACTCTTCGCGCAGAGGATACCCACCGGCGGCTAACCCCTTACC	832
Db	404	CAACAGCCTCTGTCTCAACTCTTCGCGCAGAGGATACCCACCGGCGGCTAACCCCTTACC	463
QY	833	ACACAGCATCGGCCCACTCTTACCAAGAGCTCCGTTTACAGATTTCTCTCCACCCAGCC	892
Db	464	ACACAGCATCGGCCCACTCTTACCAAGAGCTCCGTTTACAGATTTCTCTCCACCCAGCC	523
QY	893	GGGGCTGGTGCCCGGCAAGACACAGGTATCTGTGCAACAGGCCCTTTGGCTGAATTT	952
Db	524	GGGGCTGGTGCCCGGCAAGACACAGGTATCTGTGCAACAGGCCCTTTGGCTGAATTT	583
QY	953	TCACCCGGCACCAACGAGATGATCATCCAAACAGGAAAGGGGCAATGTTTCTTT	1008
Db	584	TCACCCGGCACCAACGAGATGATCATCCAAACAGGAAAGGGGCAATGTTTCTTT	639

RESULT 14					
BI917781					
LOCUS	BI917781	740 bp	mRNA	linear	EST 17-OCT-2001
DEFINITION	60318388.fpt1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5247669 5',				
	mRNA sequence.				

ORGANISM	Homo sapiens
SOURCE	Homo sapiens (human)
KEYWORDS	EST.
VERSION	BI917781.1
ACCESSION	GI:16199720
BI917781	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eumetazoa; Euarchontoglires; Primates; Catenarii;
Hominoidea; Homo.
1 (bases 1 to 740)
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: WGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

```

http://image.llnl.gov      row: b  column: 22
Plate: LRAM11625
High quality sequence stop: 740.
                        Location/Qualifiers
FEATURES
    source              1..740

```

```
/organism="Homo sapiens"
```

```

/mool_type="tRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5247669"
/lab_host="DH10B"
/clone_1b="NH_MGC_121"
/name="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned. EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC library."
```

ORIGIN

Query Match	16.7%	Score 531	DB 3	Length 740
Best Local Similarity	100.0%	Pred. No. 2.5e-268		
Matches 531	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY	529	GGAGAGCTCAGCAAGAAAGTAAACTCTCTCTGCTTTGGAGAGGGAGTCTTGAGCTTGTCACAA	588
Db	153	GGAGAGCTCAGAGAAAGTAAACTCTCTCTGCTTTGGAGAGGGAGTCTTGAGCTTGTCACAA	212
OY	589	GTTTCGATGAGCTCTGCTGACAGTACGCTACTCTCTCTCAGTTCGACGCGACCAAGTCTG	648
Db	213	GTTTCGATGAGCTCTGCTGACAGTACGCTACTCTCTCTCAGTTCGACGCGACCAAGTCTG	272
OY	649	CGGCCACTGCTCCCAAGTGCATGTTCCCGTACCCCGGCGACGACGAGACGGCGCAACCCCG	708
Db	273	CGGCCACTGCTCCCAAGTGCATGTTCCCGTACCCCGGCGACGACGAGACGGCGCAACCCCG	332
OY	709	CCTTTCATCGGAGAGCCCTTACGCGCTACATAGGCCACCAACCCGGTATACCAAGGAG	768
Db	333	CCTTTCATCGGAGAGCCCTTACGCGCTACATAGGCCACCAACCCGGTATACCAAGGAG	392
OY	769	CCTAACAAGCCTCCTGTCCAACTCTGCGCGAGAGATACCCACGGCGAGTACCCCT	828
Db	393	CCTAACAAGCCTCCTGTCCAACTCTGCGCGAGAGATACCCACGGCGAGTACCCCT	452
OY	829	ACCACAGCAGTACGGGCACTCTTACCAAGAGCTCCGTTTACCAGTTCTCTTCAACC	888
Db	453	ACCACAGCAGTACGGGCACTCTTACCAAGAGCTCCGTTTACCAGTTCTCTTCAACC	512
OY	889	AGCCGGGCGTGGTGCCTCGGCAAAAGCACAGTGTACTGTGCACAAGGCCCTTTGGCTGA	948
Db	513	AGCCGGGCGTGGTGCCTCGGCAAAAGCACAGTGTACTGTGCACAAGGCCCTTTGGCTGA	572
OY	949	AATTTACCGGCAACCAACGAGATGATCATACCAAAAGGGAAGGGGCAATGTTTCTT	1008
Db	573	AATTTACCGGCAACCAACGAGATGATCATACCAAAAGGGAAGGGGCAATGTTTCTT	632
OY	1009	TTTTAGTTTTAACATTTCTGGTCTCGATCCACGAGGCTCATTAACAATTT	1059
Db	633	TTTTAGTTTTAACATTTCTGGTCTCGATCCACGAGGCTCATTAACAATTT	683

RESULT 15	AM003023/c	LOCUS	AM003023	831 bp	mRNA	linear	EST 08-MAR-2000
DEFINITION	wg61d04.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:241571.3						
	similar to SW-TRB1 HUMAN Q16650 T-BR1N-1 PROTEIN ; contains PTR5. c3						
	TR1 repetitive element ;, mRNA sequence.						

Oy 1844 CTCTTACGGGCGGGCGGTGAAGCGCGCTGCCCTGCAGGCTGCAAGCTGCACTGGCG 1903
Db 136 CTCTTACGGGCGGGCGGTGAAGCGCGCTGCCCTGCAGGCTGCAAGCTGCACTGGCG 77
Oy 1904 CCGGCTCGGCTACTACGCCGACCCGTCGGGCTGG 1937
Db 76 CCGGCTCGGCTACTACGCCGACCCGTCGGGCTGG 43

Search completed: December 21, 2005, 01:04:47
Job time : 12035 secs

Db 241 GACTCAGCAGGGGAGCTCCAGAGAAATACCTCTCTCTGTGTGAGCGGGGTCTCTAG 300
OY 579 CTTTCGTACAGTTTCGATGGCTCTGTGTGAGATCCGTACCTCTCTCTCACTCAGCCAG 638
Db 301 CTTTCGTACAGTTTCGATGGCTCTGTGTGAGATCCGTACCTCTCTCTCACTCAGCCAG 360
OY 639 CCAAGTCGCGGCGCACTGTCCAGATGCGATGTTCCCGTACCCCGGCGAGACGAGACG 698
Db 361 CCAAGTCGCGGCGCACTGTCCAGATGCGATGTTCCCGTACCCCGGCGAGACGAGACG 420
OY 699 GCGACCCCGCTTCTCATTCGAGAGCCCTAGCCGTATCATGAGCCACACCCGGTCAATC 758
Db 421 GCGACCCCGCTTCTCATTCGAGAGCCCTAGCCGTATCATGAGCCACACCCGGTCAATC 480
OY 759 ACCAAGGAGCCTACACAGCCTCTGTCCAACTCTGTCCGCGAGGAGATACCCAGGCG 818
Db 481 ACCAAGGAGCCTACACAGCCTCTGTCCAACTCTGTCCGCGAGGAGATACCCAGGCG 540
OY 819 GCGTACCCCTACCCAGAGATAGCGGCACTCTACAGAGAGCTCCGTACAGATTC 878
Db 541 GCGTACCCCTACCCAGAGATAGCGGCACTCTACAGAGAGCTCCGTACAGATTC 600
OY 879 TCCTCACCACGAGCCGGGCTGTGTGCGCGAAAGCAGGTGTACTGTGCAACAGGCGC 938
Db 601 TCCTCACCACGAGCCGGGCTGTGTGCGCGAAAGCAGGTGTACTGTGCAACAGGCGC 660
OY 939 CTTTGGCTGAAATTTTCACCGGCAACCAAGAGATGATCATCAACAAAGGAGGCGC 998
Db 661 CTTTGGCTGAAATTTTCACCGGCAACCAAGAGATGATCATCAACAAAGGAGGCGC 720
OY 999 ATGTTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1058
Db 721 ATGTTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 780
OY 1059 TTTTGTGATGATTTTGTGCGGATCCCATCACTGAGGTTTCAAGAGGCAATGAGTT 1118
Db 781 TTTTGTGATGATTTTGTGCGGATCCCATCACTGAGGTTTCAAGAGGCAATGAGTT 840
OY 1119 CCTTGGCGAAAGCGGACCAATGTGCAAG---GAAATCGGGTCTATATGATTCGGA 1174
Db 841 CCTTGGCGAAAGCGGACCAATGTGCAAG---GAAATCGGGTCTATATGATTCGGA 900
OY 1175 TTTCCCGCAACACTGGGGCTCACTGATGCGGCAAGAAATCTCTTTTGGAAAAATTAAACT 1234
Db 901 TTTCCCGCAACACTGGGGCTCACTGATGCGGCAAGAAATCTCTTTTGGAAAAATTAAACT 960
OY 1235 TAGAACAACAAGAGCTTCAATTAACAATGAGGAGATGATGTTTCAAGTCTTGA 1294
Db 961 TAGAACAACAAGAGGCTTCAATTAACAATGAGGAGATGATGTTTCAAGTCTTGA 1020
OY 1295 CAAATACAGCCCGCTGTGATGTGTAAGTGAACAGAGACGCGACGAGAGACATAG 1354
Db 1021 CAAATACAGCCCGCTGTGATGTGTAAGTGAACAGAGACGCGACGAGAGACATAG 1080
OY 1355 CCAAGCCCGCGGTCAGACGTTCACTTCCCTGAGACTCAAGTTATGCGGTACCGC 1414
Db 1081 CCAAGCCCGCGGTCAGACGTTCACTTCCCTGAGACTCAAGTTATGCGGTACCGC 1140
OY 1415 CTACAGAAACAGGATATTACAACTGAAATAGATCAAACTTTTGGAAAAAGATT 1474
Db 1141 CTACAGAAACAGGATATTACAACTGAAATAGATCAAACTTTTGGAAAAAGATT 1200
OY 1475 TCGGATTAATTATGACAGATCTACACGCGCTGTGACATGACCGGCTGACCCCTGCGC 1534
Db 1201 TCGGATTAATTATGACAGATCTACACGCGCTGTGACATGACCGGCTGACCCCTGCGC 1260
OY 1535 CAAAGACTGCGCGCGCTGTGAGATGTGTGCGCGGCGCGCTAGACCATAGCGGCTCTTT 1594
Db 1261 CAAAGACTGCGCGCGCTGTGAGATGTGTGCGCGGCGCGCTAGACCATAGCGGCTCTTT 1320
OY 1595 CCGTGAAGACAGTGTGTGAGCAACTAGCGCAAGGCGCGCTTCAACCGGCGCGCGGCGCGC 1654
Db 1321 CCGTGAAGACAGTGTGTGAGCAACTAGCGCAAGGCGCGCTTCAACCGGCGCGCGGCGCGC 1380

OY 1655 GGGCCCCGGGCGGTTAAGGACCGGACGTGCGGCAACCAACGGGGTCGTGCGCGCA 1714
Db 1381 GGGCCCCGGGCGGTTAAGGACCGGACGTGCGGCAACCAACGGGGTCGTGCGCGCA 1440
OY 1715 GCAAGCCGAGAGACCGGGCGCGCCTTGTGCGCAACGCTGTGTTGTGACGCGCGCAACAA 1774
Db 1441 GCAAGCCGAGAGACCGGGCGCGCCTTGTGCGCAACGCTGTGTTGTGACGCGCGCAACAA 1500
OY 1775 CCGGCTGAGACTTTCGCGGCTTGTGCTATGACAGGCGCAAGACTTTCGCGGCAACGCGCGC 1834
Db 1501 CCGGCTGAGACTTTCGCGGCTTGTGCTATGACAGGCGCAAGACTTTCGCGGCAACGCGCGC 1560
OY 1835 CACGCTGTCTCTTAACGGGCGGCGGAGTGAAGGCGCTGCGTGAAGCTGACAGGCTG 1894
Db 1561 CACGCTGTCTCTTAACGGGCGGCGGAGTGAAGGCGCTGCGTGAAGCTGACAGGCTG 1620
OY 1895 CACTGAGCGCCGCTGCGCTACTAGCGGACCCGTGCGGCTGCGGCGCGCAAGTCCCGC 1954
Db 1621 CACTGAGCGCCGCTGCGCTACTAGCGGACCCGTGCGGCTGCGGCGCGCAAGTCCCGC 1680
OY 1955 GCAAGTCTGCGGCAACAAATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2014
Db 1681 GCAAGTCTGCGGCAACAAATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
OY 2015 GCGCGCGGCAATGCGCGCGCAATTCCTTACCTGCGGAGAGAGAGCGAGGCGCTGCGCGC 2074
Db 1741 GCGCGCGGCAATGCGCGCGCAATTCCTTACCTGCGGAGAGAGAGCGAGGCGCTGCGCGC 1800
OY 2075 GAGAGGCTGCGCGGCTGCGCGCGCGCGCGCGCGAGAGAGCGCAAGGCGCAAGTCTGCGA 2134
Db 1801 GAGAGGCTGCGCGGCTGCGCGCGCGCGCGCGCGAGAGAGCGCAAGGCGCAAGTCTGCGA 1860
OY 2135 TTTCAAGTGAATGAGAGCGCTCTCTGATCAATGATGATGATGATGATGATGATGATGAT 2194
Db 1861 TTTCAAGTGAATGAGAGCGCTCTCTGATCAATGATGATGATGATGATGATGATGATGAT 1920
OY 2195 TTAAGAGAGGCGCAAGCGGAGCGGAGTGTGCGCGCGCGAGAGCGCGCGTCCGAGAGTTC 2254
Db 1921 TTAAGAGAGGCGCAAGCGGAGCGGAGTGTGCGCGCGCGAGAGCGCGCGTCCGAGAGTTC 1980
OY 2255 GTCCCGGCTCAAGAGCGAGGTGTGCGGCGCAGCGGAGCTGCGAGAGAACTGCGCAAGGA 2314
Db 1981 GTCCCGGCTCAAGAGCGAGGTGTGCGGCGCAGCGGAGCTGCGAGAGAACTGCGCAAGGA 2040
OY 2315 CATTAAGCGGCTACTATGCTTCTACTGTGCAAGCTAGGCGCGCTGCGCGCGCGCGCGC 2374
Db 2041 CATTAAGCGGCTACTATGCTTCTACTGTGCAAGCTAGGCGCGCTGCGCGCGCGCGCGC 2100
OY 2375 GCGCGGCGCGGACCCCGAGCGAGCCCTGACAGCTCTTCCGAGCTCCGCTCCGACA 2434
Db 2101 GCGCGGCGCGGACCCCGAGCGAGCCCTGACAGCTCTTCCGAGCTCCGCTCCGACA 2160
OY 2435 CTCCTCTTGTGCAACCACTCAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2494
Db 2161 CTCCTCTTGTGCAACCACTCAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2220
OY 2495 TGCAGTCTCCGAGCGTATTTAACTTTTTCACAGAGCTCTGCAATTAAGCTCAC 2554
Db 2221 TGCAGTCTCCGAGCGTATTTAACTTTTTCACAGAGCTCTGCAATTAAGCTCAC 2280
OY 2555 CCACTTCAACTTTGCTGTAACTTTTGTGTTTCTGACTTCACTTCTTCTGAGAGTT 2614
Db 2281 CCACTTCAACTTTGCTGTAACTTTTGTGTTTCTGACTTCACTTCTTCTGAGAGTT 2340
OY 2615 ATCTCTTAACAATTCCTTCCCTGCTTCTTCTTCACTCTCACTTCTTCTTCTTCTTCTTCT 2674
Db 2341 ATCTCTTAACAATTCCTTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
OY 2675 ATGAAACTTTCACTTTAGAGACCTGAGGAGTCTGTGAGGAGAGAGGATTCGAGC 2734
Db 2401 ATGAAACTTTCACTTTAGAGACCTGAGGAGTCTGTGAGGAGAGAGGATTCGAGC 2460

Qy	Db	Qy	Db
491	305	551	365
AGATACAGACAAATTTTCCTCGACCTCCAGAGACTCACAGGGGACGTCAGAGAAATAACT	AGATACAGACAAATTTTCCTCGACCTCCAGAGACTCACAGGGGACGTCAGAGAAATAACT	CTCTCCCTGTTTGGACGGGGTCTC	CTCTCCCTGTTTGGACGGGGTCTC
550	364	574	388
...

RESULT 5
US-09-513-999C-22243

Sequence 22243, Application US/0951399C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Ducleit, A.Y.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
Patent No. 6783961

```

1 FILE REFERENCE: 59_US2.IMG
2 CURRENT APPLICATION NUMBER: US/09/513,999G
3 CURRENT FILING DATE: 2000-02-24
4 PRIOR APPLICATION NUMBER: US 60/122,487
5 PRIOR FILING DATE: 1999-02-26
6 NUMBER OF SEQ ID NOS: 36681
7 SOFTWARE: Patent.pm
8 SEQ ID NO 22243
9 LENGTH: 235
10 TYPE: DNA
11 ORGANISM: Homo sapiens
12 US-09-513-999G-22243

```

Query Match	7.4%	Score 235;	DB 3;	Length 235;
Best Local Similarity	100.0%;	Pred. No. 3.4e-101,		
Matches 235;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 2923 TTTTATCTCGTCAAGGACAAAACAGTTCATGCTTAACCTTTTTCCTTCTTCT 298
1 TTTTATCTCGTCAAGGACAAAACAGTTCATGCTTAACCTTTTTCCTTCTTCT 60
Db

QY	2983	TTGCTTTTCTTCTCTCATCTTCTCTCTCTCTTAAATTTCTTGAGA	304
Db	61	TTGCTTTTCTTCTCTCATCTTCTCTCTCTTAAATTTCTTGAGA	120

QY	3043	TAATATTCTAAGAGCTCTAGAAACATGAATATCTAGTAGTGATGGGTTTCCCACTTCT	3100
QY	121	TAATATTCTAAGAGCTCTAATAACATGAATACTAGTAGTAGTGAGTTTCCCACTTCT	180
Db			

QY 3103 CCTCAATCCGTTGCATGAATAATTACTATGTCCTTAATGCACACAAATAGCTA 3157
|||||
181 CCTCAATCCGTTGCATGAATAATTACTATGTCCTTAATGCACACAAATAGCTA 235
Db

RESULT 6
US-09-949-016-170893
; Sequence 170893, Application US/09949016

/ Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016

```

? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 170893

```

```

; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-170893

```

Query Match	2.0%;	Score 63;	DB 3;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 1.7e-19;		
Matches 63;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1431 ATTACACAAGCTGAAAATAGATGACCAACCCTTTGGCAAAGATTCCGGATAATATGAC 1490
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 28 ATTACACAAGCTGAAAATAGATGACCAACCCTTTGGCAAAGATTCCGGATAATATGAC 87-

Qy	1491	ACG	1493
Db	88	ACG	90

RESULT 7
US-09-513-999C-23776
; Sequence 23776, Application US/09513999C

```

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; Patent No. 6783961

```

```

1 FILE REFERENCE: 59_US-REG
2 CURRENT APPLICATION NUMBER: US/09/513,999/02
3 CURRENT FILING DATE: 2000-02-24
4 PRIOR APPLICATION NUMBER: US 60/122,487
5 PRIOR FILING DATE: 1999-02-26
6 NUMBER OF SEQ ID NOS: 36681
7 SOFTWARE: Patent.pm
8 SEQ ID NO 23776

```

```

?      LENGTH: 60
?      TYPE: DNA
?      ORGANISM: Homo sapiens
?      FEATURE:
?      NAME/KEY: misc_feature
?      LOCATION: 14
?      OTHER INFORMATION: ha a o r c o r t
?      FEATURE:
?      NAME/KEY: misc_feature
?      LOCATION: 42
?      OTHER INFORMATION: m a o r c
US-09-513-999C-23776

```

Query Match	0.8%	Score 27;	DB 3;	Length 60;
Best Local Similarity	100.0%	Pred. No. 0.022;		
Matches	27;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0.

QY 124 ATAGTCAATTGGCCAATCTTCTC 150
15 ATAGTCAATTGGCCAATCTTCTC 41

RESULT 8
US-09-949-016-13722

; Patient No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al. IN KNOW COUNCIL ASSOCIATED

1 TITLE OF INVENTION: POLYMEROPHANS 18
2
3 TITLE OF INVENTION: WITH HUMAN DISSE
4 FILE REFERENCE: CLO01307
5 CURRENT APPLICATION NUMBER: US/09/99-04-14
6 PRIORITY FILING DATE: 2000-04-14
7 PRIORITY APPLICATION NUMBER: 60/241,755
8
9 PRIORITY FILING DATE: 2000-10-20
10 PRIORITY APPLICATION NUMBER: 60/237,766
11
12 PRIORITY FILING DATE: 2000-10-03

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13722
; LENGTH: 31469
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13722

Query Match          0.8%; Score 27; DB 3; Length 31469;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2816 GAGTGGATGGATGGATGGATGGT 2842
DB      19253 GAGTGGATGGATGGATGGATGGT 19279

RESULT 9
US-09-949-016-13646/c
; Sequence 13646, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 23640
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(23640)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13646

Query Match          0.8%; Score 26; DB 3; Length 23640;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2816 GAGTGGATGGATGGATGGATGG 2841
DB      7919 GAGTGGATGGATGGATGGATGG 7894

RESULT 10
US-09-949-016-16389
; Sequence 16389, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16389
; LENGTH: 26314
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16389

Query Match          0.8%; Score 26; DB 3; Length 26314;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2816 GAGTGGATGGATGGATGGATGG 2841
DB      5513 GAGTGGATGGATGGATGGATGG 5538

RESULT 11
US-09-949-016-13006
; Sequence 13006, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13006
; LENGTH: 41243
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13006

Query Match          0.8%; Score 26; DB 3; Length 41243;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2816 GAGTGGATGGATGGATGGATGG 2841
DB      150 GAGTGGATGGATGGATGGATGG 175

RESULT 12
US-09-949-016-13424/c
; Sequence 13424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13424
; LENGTH: 85850
```

TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(85850)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13424

Query Match 0.8%; Score 26; DB 3; Length 85850;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGGATGGATGGATGGATGG 2841
DB 5916 GAGTGGATGGATGGATGGATGG 5891

RESULT 13
US-09-949-016-12362
Sequence 12362, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12362
LENGTH: 85912
TYPE: DNA
ORGANISM: Human
US-09-949-016-12362

Query Match 0.8%; Score 26; DB 3; Length 85912;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGGATGGATGGATGGATGGATGG 2844
DB 20265 TGGATGGATGGATGGATGGATGG 20290

RESULT 14
US-09-949-016-16109
Sequence 16109, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16109
LENGTH: 85913
TYPE: DNA

ORGANISM: Human
US-09-949-016-16109

Query Match 0.8%; Score 26; DB 3; Length 85913;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGGATGGATGGATGGATGGATGG 2844
DB 20265 TGGATGGATGGATGGATGGATGG 20290

RESULT 15
US-09-949-016-150203/C
Sequence 150203, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150203
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150203

Query Match 0.8%; Score 25; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGGATGGATGGATGGATGGATGG 2843
DB 185 TGGATGGATGGATGGATGGATGG 161

Search completed: December 21, 2005, 01:13:54
Job time : 545 secs

This Page Blank (uspto)

Db 61 CAGAGCCGTTTCATCGGCTGCACAAAGACAAAGATCAAAAGTGAAGCTTTTCTGATTGC 120
Qy 121 TGCATAGTGTCAATGGCCCAATCTCTTCCCAAGGAAAAAAGTAAATCAACCTT 180
Db 121 TGCATAGTGTCAATGGCCCAATCTCTTCCCAAGGAAAAAAGTAAATCAACCTT 180
Qy 181 TGAGAGCATTTGTGCTGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
Db 181 TGAGAGCATTTGTGCTGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
Qy 241 TTTATGATTAATAGGACTTTTAAAAACAGAGGACGGAGGGCGAGTTCAGTTCTAGAG 300
Db 241 TTTATGATTAATAGGACTTTTAAAAACAGAGGACGGAGGGCGAGTTCAGTTCTAGAG 300
Qy 301 CTATGAGGTGAGGACGTGCTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCATAG 360
Db 301 CTATGAGGTGAGGACGTGCTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCATAG 360
Qy 361 TGAGGACAGCTACCCACATTCAGGCGGATCCGAGCTTGTCTTGGACGATCTCCATTA 420
Db 361 TGAGGACAGCTACCCACATTCAGGCGGATCCGAGCTTGTCTTGGACGATCTCCATTA 420
Qy 421 TCTCGACCACTGACAACTCTGAGAGAGTTCACTTTGAAAAAAATTAACAAGGGAGTGA 480
Db 421 TCTCGACCACTGACAACTCTGAGAGAGTTCACTTTGAAAAAAATTAACAAGGGAGTGA 480
Qy 481 CGAATCAGTCAGATCAGACAAATTTTCTGACTCCCAAGGACTCACAGGGGACGTCCAGA 540
Db 481 CGAATCAGTCAGATCAGACAAATTTTCTGACTCCCAAGGACTCACAGGGGACGTCCAGA 540
Qy 541 GAAAGTAACTCTCTCTGTCTTGAACGGGGTCTGTGAGTCTGTGACAGTTTCTGATGCT 600
Db 541 GAAAGTAACTCTCTCTGTCTTGAACGGGGTCTGTGAGTCTGTGACAGTTTCTGATGCT 600
Qy 601 CTGCTGCGAGATGCTACTCTCTCTCTCACTCAAGCCAGCCACAGTCTGTGGCTCTCTC 660
Db 601 CTGCTGCGAGATGCTACTCTCTCTCTCACTCAAGCCAGCCAGTCTGTGGCTCTCTC 660
Qy 661 CGAGTGCACATTTCCCGTACCCCGGACAGGACCGGCGGACCCCGCTTCTCATAG 720
Db 661 CGAGTGCACATTTCCCGTACCCCGGACAGGACCGGCGGACCCCGCTTCTCATAG 720
Qy 721 GCAAGCCCTAGCCGCTACATGAGCCACCAACCCGCTCATCAACAGAGCTTACAAAGCC 780
Db 721 GCAAGCCCTAGCCGCTACATGAGCCACCAACCCGCTCATCAACAGAGCTTACAAAGCC 780
Qy 781 TCTCTGCTCACTCTCTGCGCGGAGATACCCACGGCGGGCTTACCCCTTACCCACAGCAT 840
Db 781 TCTCTGCTCACTCTCTGCGCGGAGATACCCACGGCGGGCTTACCCCTTACCCACAGCAT 840
Qy 841 ACGGCACTCTACCAAGGAGCTCGGTTCTACAGTTCTCTCAACCCAGCGCGGGAGCTGG 900
Db 841 ACGGCACTCTCTACCAAGGAGCTCGGTTCTACAGTTCTCTCAACCCAGCGCGGGAGCTGG 900
Qy 901 TGCCCGGCAAGACAGAGTGTACTGTGCAACAGGCCCCCTTGTGCTGAATTTCAACCGAC 960
Db 901 TGCCCGGCAAGACAGAGTGTACTGTGCAACAGGCCCCCTTGTGCTGAATTTCAACCGAC 960
Qy 961 ACCAAACGAGATGATCATCACCAACAGGGAAGGCGCATGTTTCTTTTAAATTTTAA 1020
Db 961 ACCAAACGAGATGATCATCACCAACAGGGAAGGCGCATGTTTCTTTTAAATTTTAA 1020
Qy 1021 ACATTTGTGCTGATGCCAGGGCTCATTAATATTTTGTGATGTGAATTTTGGGGG 1080
Db 1021 ACATTTGTGCTGATGCCAGGGCTCATTAATATTTTGTGATGTGAATTTTGGGGG 1080
Qy 1081 ATCCCAATCACTGAGGTTTCAAGAGGCAAAATGGGTTCTTGGCGAAAGCGGACACCA 1140
Db 1081 ATCCCAATCACTGAGGTTTCAAGAGGCAAAATGGGTTCTTGGCGAAAGCGGACACCA 1140
Qy 1141 ATGTGCAAGAAATGGGCTTATATGATCCGGAATTTCCCAACACTGGGGCTCACTGGA 1200
Db 1141 ATGTGCAAGAAATGGGCTTATATGATCCGGAATTTCCCAACACTGGGGCTCACTGGA 1200

Qy 1201 TGCGCAAGAAATCTTTTGGAAAAATTAACCTTACGAACAAAGAGACCTTCAATA 1260
Db 1201 TGCGCAAGAAATCTTTTGGAAAAATTAACCTTACGAACAAAGAGACCTTCAATA 1260
Qy 1261 ACAATGGGACAGATGGTGTTTTACAGTCTTGGCAAGATACAGACCCCGCTGCATGTGG 1320
Db 1261 ACAATGGGACAGATGGTGTTTTACAGTCTTGGCAAGATACAGACCCCGCTGCATGTGG 1320
Qy 1321 TGGAAATGAAAGAGACGGGACCGAGGACACTTACCCAGCCCGGCTGAGACGTTTCA 1380
Db 1321 TGGAAATGAAAGAGACGGGACCGAGGACACTTACCCAGCCCGGCTGAGACGTTTCA 1380
Qy 1381 CTTTCCCTGAGACTCAGTTCATCCGCGCTCACCGGCTACAGAACAGGATTTTACACAC 1440
Db 1381 CTTTCCCTGAGACTCAGTTCATCCGCGCTCACCGGCTACAGAACAGGATTTTACACAC 1440
Qy 1441 TGAATAATGATCACACCCCTTTTGCAAAAGGATTTGGGATTAATTAATGACAGATCTTCA 1500
Db 1441 TGAATAATGATCACACCCCTTTTGCAAAAGGATTTGGGATTAATTAATGACAGATCTTCA 1500
Qy 1501 CCGGCTGTGATGATGACCGGCTGACCCCTTGGCCCAACGATCGCGCGCTGTGAGATCG 1560
Db 1501 CCGGCTGTGATGATGACCGGCTGACCCCTTGGCCCAACGATCGCGCGCTGTGAGATCG 1560
Qy 1561 TGCCCGGGGCGCGGCTACGCGATGGCCGAGCTCTTCTGTGAGAGACAGTTTCTGAGCACT 1620
Db 1561 TGCCCGGGGCGCGGCTACGCGATGGCCGAGCTCTTCTGTGAGAGACAGTTTCTGAGCACT 1620
Qy 1621 ACGCCAAAGCCGCTTCCACCCGGGCGCGGGCGCGGCGCGGATACGGAACCGCA 1680
Db 1621 ACGCCAAAGCCGCTTCCACCCGGGCGCGGGCGCGGCGCGGATACGGAACCGCA 1680
Qy 1681 GCGTGCAGCAACCAAGGGGTGCTGTCCCGGACAGGCGGAGAGACCCGGGCGCGGCT 1740
Db 1681 GCGTGCAGCAACCAAGGGGTGCTGTCCCGGACAGGCGGAGAGACCCGGGCGCGGCT 1740
Qy 1741 CGCGGCAACCGTGGTTTGTGACGCGCGGCAACACCGGCTTGAATTGCGGCGCTCGGCT 1800
Db 1741 CGCGGCAACCGTGGTTTGTGACGCGCGGCAACACCGGCTTGAATTGCGGCGCTCGGCT 1800
Qy 1801 ATGACACGGCCACGGAATTCGCGGGCAACCGGACGCTGTCTTTACCGCGGCGGCG 1860
Db 1801 ATGACACGGCCACGGAATTCGCGGGCAACCGGACGCTGTCTTTACCGCGGCGGCG 1860
Qy 1861 GCGTGAAGGCGCTCGCTGAGGCTGAGGCTGACAGTCTGAGGCGGCGCTCGGCTACG 1920
Db 1861 GCGTGAAGGCGCTCGCTGAGGCTGAGGCTGACAGTCTGAGGCGGCGCTCGGCTACG 1920
Qy 1921 CCGACCCGTCGGGCTGGGGGCGCCGACAGTCCCGGCAATCTGCGGACCAAGTCGAGGCT 1980
Db 1921 CCGACCCGTCGGGCTGGGGGCGCCGACAGTCCCGGCAATCTGCGGACCAAGTCGAGGCT 1980
Qy 1981 CGGTGTGCTGTGTGCGCCCAACAGCGCGCGGCGCGCGGATGAGTGGCGCGCAATC 2040
Db 1981 CGGTGTGCTGTGTGCGCCCAACAGCGCGCGGCGCGCGGATGAGTGGCGCGCAATC 2040
Qy 2041 CTTACTTGGGGAAGAGGCGGAGGCTGTGGCGGCGGAGCGCTTCCGCTGTGCGGCGGCG 2100
Db 2041 CTTACTTGGGGAAGAGGCGGAGGCTGTGGCGGCGGAGCGCTTCCGCTGTGCGGCGGCG 2100
Qy 2101 CCGCGGAGAGCGCCAAAGCCCAAGGACTGTCCGATTCAGTGTGATGAGACGCGCTCT 2160
Db 2101 CCGCGGAGAGCGCCAAAGCCCAAGGACTGTGTCCGATTCAGTGTGATGAGACGCGCTCT 2160
Qy 2161 CGATCAAGTCCATGACTTCCAGCACTCGGAGATTTTACGACAGGCGCAAGCGAGCGGA 2220
Db 2161 CGATCAAGTCCATGACTTCCAGCACTCGGAGATTTTACGACAGGCGCAAGCGAGCGGA 2220
Qy 2221 TCTGCGCGGCGGACAGCGCGGCTGTCCGAGATTTGTCTCCGCTTCAAGAGCGAGGTGCTGG 2280
Db 2221 TCTGCGCGGCGGACAGCGCGGCTGTCCGAGATTTGTCTCCGCTTCAAGAGCGAGGTGCTGG 2280

QY 2281 CCCACGGGACCTGCGAAGAACTGCGCCAGACATTAGCGGCTTACTATGCGCTTCT 2340
 Db 2281 CCCACGGGACCTGCGAAGAACTGCGCCAGACATTAGCGGCTTACTATGCGCTTCT 2340
 QY 2341 CGCAGACGCTGAGCGCGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2400
 Db 2341 CGCAGACGCTGAGCGCGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2400
 QY 2401 CCTCAAGCTCTTCCCGACGCTCGCGCTCCGACACCTCTCTTGGCGACCCACTATTT 2460
 Db 2401 CCTCAAGCTCTTCCCGACGCTCGCGCTCCGACACCTCTCTTGGCGACCCACTATTT 2460
 QY 2461 ATTGACCTCGATGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2520
 Db 2461 ATTGACCTCGATGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2520
 QY 2521 CTTTTTGGACAGCGCTCTGCAATTTAGCTACCGACCTTCACTTTGCTTAAACCTT 2580
 Db 2521 CTTTTTGGACAGCGCTCTGCAATTTAGCTACCGACCTTCACTTTGCTTAAACCTT 2580
 QY 2581 TTGGTTTCTCTACTTACTCTTCTTCTGAGAGTTATCTCTAACAATTCCTCCCTCC 2640
 Db 2581 TTGGTTTCTCTACTTACTCTTCTTCTGAGAGTTATCTCTAACAATTCCTCCCTCC 2640
 QY 2641 GTCTTTCTCTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2700
 Db 2641 GTCTTTCTCTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2700
 QY 2701 TGGGCGAGCTGCTGCGAGCGAGCGAGCTGCGAGCGAGCTGCGAGCGAGCTGCGAGC 2760
 Db 2701 TGGGCGAGCTGCTGCGAGCGAGCGAGCTGCGAGCGAGCTGCGAGCGAGCTGCGAGC 2760
 QY 2761 CATAGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTTATCCCGAGTG 2820
 Db 2761 CATAGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTTATCCCGAGTG 2820
 QY 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
 Db 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
 QY 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940
 Db 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940
 QY 2941 ACAAAACGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTTCTCTCC 3000
 Db 2941 ACAAAACGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTCTCTCC 3000
 QY 3001 TCTCATACTTCT 3060
 Db 3001 TCTCATACTTCT 3060
 QY 3061 TAAACACGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTCTCTCTCT 3120
 Db 3061 TAAACACGAGTGTGACTGTAGAACCTGAGCCACCGAGCGGCTCTTCTCTCTCTCT 3120
 QY 3121 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3180
 Db 3121 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3180
 QY 3181 TAAAGG 3186
 Db 3181 TAAAGG 3186

RESULT 2
 US-10-029-386-20697
 ; Sequence 20697, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G;
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: AROMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029, 386
 ; NUMBER OF SEQ ID NOS: 3428
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 20697
 ; LENGTH: 864
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC009487.3
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
 ; OTHER INFORMATION: NT HIT: g1142971, EVALUATE 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: Q1650, EVALUATE 0.00e+00
 ; OTHER INFORMATION: EST_HUMAN HIT: AI201654.1, EVALUATE 0.00e+00
 ; US-10-029-386-20697
 Query Match 27.1%; Score 864; DB 6; Length 864;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1493 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAGACGACTGCGCGCTC 1552
 Db 1 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAGACGACTGCGCGCTC 60
 QY 1553 GCAAGTCTGCG 1612
 Db 61 GCAAGTCTGCG 120
 QY 1613 GAGCACTACGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1672
 Db 121 GAGCACTACGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 1673 GAGCGAGCG 1732
 Db 181 GAGCGAGCG 240
 QY 1733 CGCGCGCTGCG 1792
 Db 241 CGCGCGCTGCG 300
 QY 1793 CTGCGGCTAGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1852
 Db 301 CTGCGGCTAGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 1853 GCG 1912
 Db 361 GCG 420
 QY 1913 CTACTACGCGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
 Db 421 CTACTACGCGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 QY 1973 GTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 2032
 Db 481 GTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 540
 QY 2033 CGCAATCCCTACTCTGCGGCGAGAGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCT 2092
 Db 541 CGCAATCCCTACTCTGCGGCGAGAGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCT 600
 QY 2093 GCCCG 2152
 Db 601 GCCCG 660
 QY 2153 GCCCTCTCGATCAAGTCCATGCACTCGAGCACTCGGAGATTAGAGAGCGCGCAAGCG 2212

Db 661 GGCCTCTGATCAAGTCCATGACTCCAGCAGCTCGGGATTTCAGAGGCCAACG 720
Qy 2213 GAGCGGATCTCGCCGCCGACACCCGCTGTCCAGAGTTGTCGCCGCTCAAGAGCGA 2272
Db 721 GAGCGGATCTCGCCGCCGACACCCGCTGTCCAGAGTTGTCGCCGCTCAAGAGCGA 780
Qy 2273 GGTGTCGCCGCGGAGTCTGCAAGAACTGCGCAAGACATTAGCGGCTACTATGG 2332
Db 781 GGTGTCGCCGCGGAGTCTGCAAGAACTGCGCAAGACATTAGCGGCTACTATGG 840
Qy 2333 CTTTACTGCGACAGCTAGGCCG 2356
Db 841 CTTTACTGCGACAGCTAGGCCG 864

RESULT 3
US-10-029-386-6986
; Sequence 6986, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6986
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009487.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: H10108.1, EVALU8 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q16650, EVALU8 5.00e-15
; OTHER INFORMATION: NT HIT: g11429971, EVALU8 0.00e+00
US-10-029-386-6986

Query Match 16.2%; Score 515; DB 6; Length 515;
Best Local Similarity 100.0%; Pred. No. 1,1e-257;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2236 CGCCCGTGTCCAGAGTTGTCGCCGCTCAAGAGCGAGGTGTCGCCGCGGAGCTGCG 2295
Db 1 CGCCCGTGTCCAGAGTTGTCGCCGCTCAAGAGCGAGGTGTCGCCGCGGAGCTGCG 60
Qy 2296 AAGAACTGCGCGCAAGACATTAGGCGCTACTATGCTTCTACTGCAAGACTAGCGCG 2355
Db 61 AAGAACTGCGCGCAAGACATTAGGCGCTACTATGCTTCTACTGCAAGACTAGCGCG 120
Qy 2356 CGCCG 2415
Db 121 CGCCG 180
Qy 2416 CGAGCTCGCGCTCCCACTCTCTCTGCGACCCACTCATTTTATTTGACCTCGATG 2475
Db 181 CGAGCTCGCGCTCCCACTCTCTCTGCGACCCACTCATTTTATTTGACCTCGATG 240
Qy 2476 GCGGCTGCGAGGAAATAGTGAAGTCTCCGAGCGTGAATTTAACCTTTTGGACAGCA 2535
Db 241 GCGGCTGCGAGGAAATAGTGAAGTCTCCGAGCGTGAATTTAACCTTTTGGACAGCA 300
Qy 2536 GTCCTGCAATTAAGCTCACGACCTTCAACTTGTGCTAAACCTTTTGGTTCGACTT 2595

Db 301 GTCCTGGAATTAACCTACCGGACCTTCACTTGTCTTAACCTTTGGTTTCTACTT 360
Qy 2596 ACTCTTCTTGTGTGAGTATCTCTCAAGATTTCCCTCCCTCGCTTTCTTACTT 2655
Db 361 ACTCTTCTTGTGTGAGTATCTCTCAAGATTTCCCTCCCTCGCTTTCTTACTT 420
Qy 2656 CCTACTTCTCTTCTTGTGAATGAACCTTCACTTTAGAGAGCTGAGAGTCTGTCA 2715
Db 421 CCTACTTCTCTTCTTGTGAATGAACCTTCACTTTAGAGAGCTGAGAGTCTGTCA 480
Qy 2716 GGCAGACGCAATTCGACCCGCGCAAGTCTGCGCT 2750
Db 481 GGCAGACGCAATTCGACCCGCGCAAGTCTGCGCT 515

RESULT 4
US-10-842-072-37/c
; Sequence 37, Application US/10842072
; Publication No. US20050158731A1
; GENERAL INFORMATION:
; APPLICANT: PLASS, CHRISTOPH
; TITLE OF INVENTION: DETECTION OF METHYLATED CpG RICH SEQUENCES DIAGNOSTIC FOR
; FILE OF INVENTION: MALIGNANT CELLS
; FILE REFERENCE: 22727-04242
; CURRENT APPLICATION NUMBER: US/10/842,072
; CURRENT FILING DATE: 2004-05-10
; PRIOR APPLICATION NUMBER: 09/775,398
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-842-072-37

Query Match 16.1%; Score 514; DB 9; Length 514;
Best Local Similarity 100.0%; Pred. No. 3.6e-257;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1504 GCTGTGATGAGACCGCTGACCCCTCGGCCCAAGACTGCGCGCTCGCAGATCGTC 1563
Db 514 GCTGTGATGAGACCGCTGACCCCTCGGCCCAAGACTGCGCGCTCGCAGATCGTC 455
Qy 1564 CCGGGGCGCGTACGCAATGCGCGCTTTCTTCTGAGAGACCAATTTGTAGCAACTAG 1623
Db 454 CCGGGGCGCGTACGCAATGCGCGCTTTCTTCTGAGAGACCAATTTGTAGCAACTAG 395
Qy 1624 CCAAGCGCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683
Db 394 CCAAGCGCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
Qy 1684 TGCAGACACCAAGGCTGTCTGCGCGAGCAAGGCGCAAGACCGCGCGCGCGCTGCG 1743
Db 334 TGCAGACACCAAGGCTGTCTGCGCGAGCGCGAGGCGCAAGACCGCGCGCGCGCTGCG 275
Qy 1744 CGCAACGCTGTTGTGACCGCGCGCGCGCGCAACCGGCTGAGATTGCGGCGCTGATG 1803
Db 274 CGCAACGCTGTTGTGACCGCGCGCGCGCGCAACCGGCTGAGATTGCGGCGCTGATG 215
Qy 1804 ACAAGGCGCAAGGACCTTGGCGGGGCAAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1863
Db 214 ACAAGGCGCAAGGACCTTGGCGGGGCAAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 155
Qy 1864 TGAAGGCGCTGCGCTGAGGCTGCAAGGCTGCACTGCGCGCGCGCTGCGCTACTACGCG 1923
Db 154 TGAAGGCGCTGCGCTGAGGCTGCAAGGCTGCACTGCGCGCGCGCTGCGCTACTACGCG 95
Qy 1924 ACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983
Db 94 ACCGCTGCGGCTGCGGCG 35

Query Match 1.0%; Score 32; DB 8; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
DB 112 ACCCCCTGCGCCACGACTGCGCGCTCGCA 143

RESULT 9
US-10-363-483A-9001/c
; Sequence 9001, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9001
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 9001
US-10-363-483A-9001

Query Match 1.0%; Score 32; DB 9; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
DB 915 ACCCCCTGCGCCACGACTGCGCGCTCGCA 864

RESULT 10
US-10-363-483A-9002
; Sequence 9002, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9002
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 9002
US-10-363-483A-9002

Query Match 1.0%; Score 32; DB 9; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
DB 112 ACCCCCTGCGCCACGACTGCGCGCTCGCA 143

RESULT 11
US-10-027-632-273001/c
; Sequence 273001, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273001
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-273001

Query Match 0.8%; Score 27; DB 5; Length 1006;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGGATGATGATGATGATGATGATG 2845
DB 894 TGGATGATGATGATGATGATGATG 868

RESULT 12
US-10-027-632-273001/c
; Sequence 273001, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273001
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-273001

Query Match 0.8%; Score 27; DB 6; Length 1006;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGATGATGATGATGATGATGATG 2845
 |||||
 DB 894 TGATGATGATGATGATGATGATG 868

RESULT 13

US-10-674-124A-16640
 ; Sequence 16640, Application US/10674124A
 ; Publication No. US2004019797A1
 ; GENERAL INFORMATION:

APPLICANT: INOKO, Hideoshi
 APPLICANT: TAMITA, Gen
 TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
 FILE REFERENCE: ORIN-003CIP

CURRENT APPLICATION NUMBER: US/10/674,124A
 CURRENT FILING DATE: 2003-09-26
 PRIOR APPLICATION NUMBER: 10/257,511
 PRIOR FILING DATE: 2003-03-07
 PRIOR APPLICATION NUMBER: PCT/JP00/07621
 PRIOR FILING DATE: 2000-10-30
 PRIOR APPLICATION NUMBER: JP2000-112699
 PRIOR FILING DATE: 2000-04-13
 PRIOR APPLICATION NUMBER: JP2002-327516
 PRIOR FILING DATE: 2002-09-28
 PRIOR APPLICATION NUMBER: JP2002-383869
 PRIOR FILING DATE: 2002-12-09
 NUMBER OF SEQ ID NOS: 27110
 SEQ ID NO 16640

LENGTH: 182
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: D20S208
 FEATURE:
 OTHER INFORMATION: Located on chromosome 10
 FEATURE:
 OTHER INFORMATION: Distance between a terminus base of telomere on
 OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
 OTHER INFORMATION: sequence : 103456393
 FEATURE:
 OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
 OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
 OTHER INFORMATION: 5'-terminus of this base sequence : 52125
 US-10-674-124A-16640

Query Match 0.8%; Score 26; DB 8; Length 182;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 |||||
 DB 107 GAGTGATGATGATGATGATG 132

RESULT 14

US-10-719-993-31636/c
 ; Sequence 31636, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001496
 CURRENT APPLICATION NUMBER: US/10/719,993
 CURRENT FILING DATE: 2003-11-24
 NUMBER OF SEQ ID NOS: 55342

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 31636
 LENGTH: 201
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-719-993-31636

Query Match 0.8%; Score 26; DB 8; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 |||||
 DB 150 GAGTGATGATGATGATGATG 125

RESULT 15

US-10-719-993-31637/c
 ; Sequence 31637, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001496
 CURRENT APPLICATION NUMBER: US/10/719,993
 CURRENT FILING DATE: 2003-11-24
 NUMBER OF SEQ ID NOS: 55342
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 31637
 LENGTH: 201
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-719-993-31637

Query Match 0.8%; Score 26; DB 8; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 |||||
 DB 149 GAGTGATGATGATGATGATG 124

Search completed: December 21, 2005, 01:54:54
 Job time : 2454 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 07:04:33 ; Search time 15679 Seconds
(without alignments)
11550.688 Million cell updates/sec

Title: US-10-649-400-3

Sequence: 1 caggtgattatcctaataa.....caccacaacacttaagg 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sta:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2884.4	90.5	2910	6	CQ729238 Sequence
2	2753	86.4	2894	8	HSU49250 Human putat
3	2490	78.2	3992	9	BC052737 Mus muscu
4	2400.4	75.3	3897	9	BC058399 Mus muscu
5	2334.8	73.3	3814	9	MMU9251 Mus muscu
6	1793	56.3	1793	6	CQ769652 Sequence
7	1793	56.3	1793	6	BC029289 Homo sapi
8	1695.4	53.2	18366	8	AC009487 Homo sapi
9	1289	40.5	101040	14	AC165777 Bos tauri
10	1164.6	36.6	209914	9	AL845291 Mouse DNA
11	1157.8	36.3	176172	14	AC127930 Rattus no
12	971	30.5	101040	14	AC165777 Bos tauri
13	694.8	21.8	2218	5	AF287006 Danio rer
14	514	16.1	514	5	AR655897 Sequence
15	491.2	15.4	793	8	HS341671 Homo sapi
16	420.4	13.2	2900	5	XUW5996 Xenopus lae
17	417.6	13.1	690	8	HS323162 Homo sapi
18	412	12.9	3038	5	AB019785 Cynops py

19	407.6	12.8	2986	5	BC084243	BC084243 Xenopus l
20	407.4	12.8	2480	5	AF329830	AF329830 Danio rer
21	407.4	12.8	2572	5	BC067719	BC067719 Danio rer
22	404.2	12.7	2450	5	AF287007	AF287007 Danio rer
23	401.6	12.6	689	8	HS342462	HS342462 Homo sapi
24	395.8	12.4	2842	5	CT025431	CT025431 Xenopus l
25	385.2	12.1	641	8	HS338878	HS338878 Homo sapi
26	384.6	12.1	388	6	BD024632	BD024632 Sequence
27	384.6	12.1	388	6	AX885022	AX885022 Sequence
28	374.6	11.8	2494	6	AR169779	AR169779 Sequence
29	374.6	11.8	2840	8	BC037568	BC037568 Homo sapi
30	374.6	11.8	3577	9	BC094319	BC094319 Mus muscu
31	356.6	11.2	784	8	HS3427080	HS3427080 Homo sapi
32	353	11.1	1894	6	CQ723583	CQ723583 Sequence
33	350.8	11.0	2883	8	AB031037	AB031037 Mus muscu
34	346	10.9	2061	8	AB031038	AB031038 Homo sapi
35	341.2	10.7	2756	6	CS035351	CS035351 Sequence
36	341.2	10.7	2756	6	CS036706	CS036706 Sequence
37	341.2	10.7	2756	6	CS044303	CS044303 Sequence
38	341.2	10.7	2756	6	CS045658	CS045658 Sequence
39	330.2	10.4	741	8	HS333254	HS333254 Homo sapi
40	328.2	10.3	691	8	HS333038	HS333038 Homo sapi
41	326.2	10.2	579	9	AY457971	AY457971 Rattus no
42	321.8	10.1	693	8	HS341518	HS341518 Homo sapi
43	321.4	10.1	540	9	AF013281	AF013281 Mus muscu
44	316.2	9.9	1608	6	AX055554	AX055554 Sequence
45	316.2	9.9	1608	6	AX695541	AX695541 Sequence

ALIGNMENTS

RESULT 1
LOCUS CQ729238 2910 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15172 from Patent WO02068579.
ACCESSION CQ729238
VERSION CQ729238.1 GI:42300290

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kite, such as nucleic acid arrays, comprising a majority of
humaneons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 15172 06-SEP-2002;
JOURNAL FE Corporation (NY) (US)
Location/Qualifiers
1..2910
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN
Query Match 90.5%; Score 2884.4; DB 6; Length 2910;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2907; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY	279	GCGAGTGTTCAGGTTCTGAGCTATGACGCTGAGACATGCTTCTCTTATCAG	338
DB	1	GCGAGTGTTCAGGTTCTGAGCTATGACGCTGAGACATGCTTCTCTTATCAG	60
QY	339	CTCTCAAGAAATTTCTCATGATGACAGACAGCTACCATTCAGCGGATCCGAGCTT	398
DB	61	CTCTCAAGAAATTTCTCATGATGACAGACAGCTACCATTCAGCGGATCCGAGCTT	120
QY	399	GTCCTGACAGATCATCCATTCATCTGACACCTGACACCTGAGAGAAAGTTGACCTTTG	458
DB	121	GTCCTGACAGATCATCCATTCATCTGACACCTGACACCTGAGAGAAAGTTGACCTTTG	180

QY 459 AAAAAAAAAATTAACGAGGGGGATGAAGATCATGATGATACAGACAATTTTCTGTGACTCCAG 518
Db 181 AAAAAAAAAATTAACGAGGGGGATGAAGATCATGATGATACAGACAATTTTCTGTGACTCCAG 240
QY 519 GACTCAGCAGGGGAGCTCCAGAGATTAACCTCTCTCTGTCTTGAAGGGGCTCTGAG 578
Db 241 GACTCAGCAGGGGAGCTCCAGAGATTAACCTCTCTCTGTCTTGAAGGGGCTCTGAG 300
QY 579 CTTTGTGACAGTTTGGATGAGCTCTGTGAGATTCGCTACTCTCTCTGATCCAGCCAG 638
Db 301 CTTTGTGACAGTTTGGATGAGCTCTGTGAGATTCGCTACTCTCTCTGATCCAGCCAG 360
QY 639 CCACAGTCTGGGGCCACTGCTCCAGTGCATGTTCCGGTACCCGGGCGAGACGGACCG 698
Db 361 CCACAGTCTGGGGCCACTGCTCCAGTGCATGTTCCGGTACCCGGGCGAGACGGACCG 420
QY 699 GCGCACCCCGGCTTCTCATCGGAGCCCTAGCCGGTACATGAGCCACACCCGGTCAATC 758
Db 421 GCGCACCCCGGCTTCTCATCGGAGCCCTAGCCGGTACATGAGCCACACCCGGTCAATC 480
QY 759 ACCAAGCAGGCTTACACAGGCTCTGTCTCAACTCTGCGCGGAGGATACCCACGGCC 818
Db 481 ACCAAGCAGGCTTACACAGGCTCTGTCTCAACTCTGCGCGGAGGATACCCACGGCC 540
QY 819 GGCTACCCCTTACCCAGAGATACGGGCACTCTCAACAGAGACTCCGTTCTACAGTTTC 878
Db 541 GGCTACCCCTTACCCAGAGATACGGGCACTCTCAACAGAGACTCCGTTCTACAGTTTC 600
QY 879 TCCTCACCACGACCGGGGCTGTGTGCTCGGCAAGCAAGTGTACTCTGTGCAAGGCCC 938
Db 601 TCCTCACCACGACCGGGGCTGTGTGCTCGGCAAGCAAGTGTACTCTGTGCAAGGCCC 660
QY 939 CTTTGGCTGAAATTTTCAACGGGCAACCAACGAGATGATATCAACCAACGAGGAGGGCC 998
Db 661 CTTTGGCTGAAATTTTCAACGGGCAACCAACGAGATGATATCAACCAACGAGGAGGGCC 720
QY 999 ATGTTTCCCTTTTAAAGTTTAAACATTTCTGGTCTCGATCCGACGGGCTCATTAACAATTT 1058
Db 721 ATGTTTCCCTTTTAAAGTTTAAACATTTCTGGTCTCGATCCGACGGGCTCATTAACAATTT 780
QY 1059 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGGTTTCAAGAGGCAATGGGTT 1118
Db 781 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGGTTTCAAGAGGCAATGGGTT 840
QY 1119 CTTTGGCGGAAAGCGGACCAATGTGCAAGGAAATCGGTTCTAATATGATCCGGAATTC 1178
Db 841 CTTTGGCGGAAAGCGGACCAATGTGCAAGGAAATCGGTTCTAATATGATCCGGAATTC 900
QY 1179 CCCAACACTGGGGCTCACTGATGCGCCAAAG-AAATCTTTTGGAAATTTAAACTTAC 1237
Db 901 CCCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGGAAATTTAAACTTAC 960
QY 1238 GAACAACAAA--GAGAGTTCAAAATAACAATGAGCAGATGATGTTTACAGTCTTGACA 1296
Db 961 GAACAACAAAAGGAGGCTTCAAAATAACAATGAGCAGATGATGTTTACAGTCTTGACA 1020
QY 1297 AGTACACGCGCCGCTGCAATGTGTGGAAGTGAAGAGACGGCACCGAGACACTAGCC 1356
Db 1021 AGTACACGCGCCGCTGCAATGTGTGGAAGTGAAGAGACGGCACCGAGAGCACTAGCC 1080
QY 1357 AGCCGGGCGGGGAGAGAGTTCACTTTCCTGAGACTGATTCATTCGCGCAACGGGCT 1416
Db 1081 AGCCGGGCGGGGAGAGAGTTCACTTTCCTGAGACTGATTCATTCGCGCAACGGGCT 1140
QY 1417 ACCAGAACGAGATATTACAACTGAAATATGATACAAACCTTTTGGAAAGAGATTC 1476
Db 1141 ACCAGAACGAGATATTACAACTGAAATATGATACAAACCTTTTGGAAAGAGATTC 1200
QY 1477 GGGATTAATTATGACAGATCTACACGGGCTGTGACATGAGACCGCTGACCCCTTGCCCA 1536
Db 1201 GGGATTAATTATGACAGATCTACACGGGCTGTGACATGAGACCGCTGACCCCTTGCCCA 1260

QY 1537 ACGACTGCGCGGCTCGAGATCGTGCCCGGGGCGCCGCTACGCCATGAGCGGCTCTTCC 1596
Db 1261 ACGACTGCGCGGCTCGAGATCGTGCCCGGGGCGCCGCTACGCCATGAGCGGCTCTTCC 1320
QY 1597 TGCAGAACAGTTCTGTAGCAACTACCGCAAGGCTCCGCTTCACTCCGGGCGCGGGCGG 1656
Db 1321 TGCAGAACAGTTCTGTAGCAACTACCGCAAGGCTCCGCTTCACTCCGGGCGCGGGCGG 1380
QY 1657 GCGCGGGGCGGGGATCGGACCGGAGGTCGCGCACCAACAGGGCTGTGTGCGCGGAGC 1716
Db 1381 GCGCGGGGCGGGGATCGGACCGGAGGTCGCGCACCAACAGGGCTGTGTGCGCGGAGC 1440
QY 1717 AGGCGAGAACCCCGGCGCGGCTCGCGCAACGCTGTGTTGTACAGCGCGGCAACACC 1776
Db 1441 AGGCGAGAACCCCGGCGCGGCTCGCGCAACGCTGTGTTGTACAGCGCGGCAACACC 1500
QY 1777 GGGTGAACCTTGGGAGCTCGGCTATGACAGGCTACAGGACTTGTGGGAGCAACGGGCA 1836
Db 1501 GGGTGAACCTTGGGAGCTCGGCTATGACAGGCTACAGGACTTGTGGGAGCAACGGGCA 1560
QY 1837 GGTGTCTCTTTAAGCGGCGGGCGGGGCTGTGAAGGGGCTGTGCGGATGACAGGCTGCA 1896
Db 1561 GGTGTCTCTTTAAGCGGCGGGCGGGGCTGTGAAGGGGCTGTGCGGATGACAGGCTGCA 1620
QY 1897 CTGGCGGCGCGCTGAGCTACTACGCGCACCGGCTGGGCTGGGAGCGCCGAGTCCCGC 1956
Db 1621 CTGGCGGCGCGCTGAGCTACTACGCGCACCGGCTGGGCTGGGAGCGCCGAGTCCCGC 1680
QY 1957 AGTACTGGGAGCAAGTGGGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2016
Db 1681 AGTACTGGGAGCAAGTGGGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1740
QY 2017 CCGGCGGATGAGCGGCGGCGCAATCCCTAACCCTGGGCGAGAGGCGGAGGCGCTGGCGCG 2076
Db 1741 CCGGCGGATGAGCGGCGGCGCAATCCCTAACCCTGGGCGAGAGGCGGAGGCGCTGGCGCG 1800
QY 2077 AGGCTGTGCGGCTGCGGCGGCGCGCGCGGCGGAGGAGCGGCAAGGCTGTGCTGTGCTGTG 2136
Db 1801 AGGCTGTGCGGCTGCGGCGGCGCGCGGCGGCGGAGGAGGCGGCAAGGCTGTGCTGTGCTGTG 1860
QY 2137 CCAAGTGTGATGAGACGCGCTCTGTGATCAAGTCAATGACTCCAGGAGCTGTGGGATTT 2196
Db 1861 CCAAGTGTGATGAGACGCGCTCTGTGATCAAGTCAATGACTCCAGGAGCTGTGGGATTT 1920
QY 2197 ACGAGAGGCGCAACGAGAGGAGGAGTCTGTGCGGCGGCAAGGCGGCTGTGCGAGGTTCT 2256
Db 1921 ACGAGAGGCGCAACGAGAGGAGGAGTCTGTGCGGCGGCAAGGCGGCTGTGCGAGGTTCT 1980
QY 2257 CCGGCTCAAGAGGAGGAGGAGTGTGCGCGGAGCTGTGAGAGAACTGTGCGCAAGACA 2316
Db 1981 CCGGCTCAAGAGGAGGAGGAGTGTGCGCGGAGCTGTGAGAGAACTGTGCGCAAGACA 2040
QY 2317 TTAAGGAGTACTATGAGCTTCTTACTGTGCAACGCTAGGCGGCTGTGCGGCGGCGCGC 2376
Db 2041 TTAAGGAGTACTATGAGCTTCTTACTGTGCAACGCTAGGCGGCTGTGCGGCGGCGCGC 2100
QY 2377 GCGGCGCGGAGCGCCGAGGCGCGGCTGTGCAAGGCTTCTGCGGCTGTGCGGCGGCGCACT 2436
Db 2101 GCGGCGCGGAGCGCCGAGGCGCGGCTGTGCAAGGCTTCTGCGGCTGTGCGGCGGCGCACT 2160
QY 2437 CCTCCTTGTGCAACCACTCATTTTATTTTGAACCTTGTGATGAGCGCTGTGAGGAAATGAG 2496
Db 2161 CCTCCTTGTGCAACCACTCATTTTATTTTGAACCTTGTGATGAGCGCTGTGAGGAAATGAG 2220
QY 2497 CAGGTCTCGAGGCTGATTTTAACTTTTGTGACACAGGCTGTGCAATTAAGTCAACG 2556
Db 2221 CAGGTCTCGAGGCTGATTTTAACTTTTGTGACACAGGCTGTGCAATTAAGTCAACG 2280
QY 2557 ACCTTCAACTTTTGTGTAACCTTTTGTGTTTCTTACTTACTCTTCTTCTGTGAGTTAT 2616
Db 2281 ACCTTCAACTTTTGTGTAACCTTTTGTGTTTCTTACTTACTCTTCTTCTGTGAGTTAT 2340
QY 2617 CCTCTCAAAATTCCTTCCCTGTGTCTTCTTCTTAACTCTTACTTCTTCTTGTATAT 2676

QY	939	CTTTGGCTGAAATTTACCCGGGCACCAAACGAGATGATCATACCAAACAGGAAAGCGC	998
Db	661	CTTTGGCTGAAATTTACCCGGGCACCAAACGAGATGATCATACCAAACAGGAAAGCGC	720
QY	999	ATGTTTCTTTTTTAAAGTTTAACTTTCGTGGTCTCGATGCCAGGGCTCATTAACAATTT	1058
Db	721	ATGTTTCTTTTTTAAAGTTTAACTTTCGTGGTCTCGATGCCAGGGCTCATTAACAATTT	780
QY	1059	TTTGTGATGTGATTTTGGCGGATCCCAATCATCTGAGGTTTCAAGAGGCAATGGGTT	1118
Db	781	TTTGTGATGTGATTTTGGCGGATCCCAATCATCTGAGGTTTCAAGAGGCAATGGGTT	840
QY	1119	CCTTGCGGCAAAAGCGGACACCAATGTGCAAGGAAATCGGGTCTAATATGATCCGATTC	1178
Db	841	CTTTCGGGCAAAAGCGGACACCAATGTGCAAGGAAATCGGGTCTAATATGATCCGATTC	900
QY	1179	CCCAACACTGGGGCTCACTGATGCGCCAGAAATCTCTTTGGAAAAATTAAACTTAAC	1238
Db	901	CCCAACACTGGGGCTCACTGATGCGCCAGAAATCTCTTTGGAAAAATTAAACTTAAC	960
QY	1239	AACACAAAGAGCTTCAATTAACAATGGGCGATGTGTGTTTCACTCTTGGACAAG	1299
Db	961	AACACAAAGAGCTTCAATTAACAATGGGCGATGTGTGTTTCACTCTTGGACAAG	1020
QY	1289	TACGAGCCCCGCTCATGTGGTGTGAAGTGAACGAGAGACGGGACAGGACACTAGCCAG	1358
Db	1021	TACGAGCCCCGCTCATGTGGTGTGAAGTGAACGAGAGACGGGACAGGACACTAGCCAG	1088
QY	1359	CCCGGCGCGGTGCAGACGTTCACTTTCCCTGAGACTCACTTATCGCGTCAACGCGCTAC	1418
Db	1081	CCCGGCGCGGTGCAGACGTTCACTTTCCCTGAGACTCACTTATCGCGTCAACGCGCTAC	1140
QY	1419	CAGAACAGGATATTACAACATGAAAAATAGATTCACACCCCTTTGGCAAAAGGATTTGG	1478
Db	1141	CAGAACAGGATATTACAACATGAAAAATAGATTCACACCCCTTTGGCAAAAGGATTTGG	1200
QY	1479	GATATATTATGACACGATCTACACCGGCTGTGACATGAGACCGGCTGACCCCTGCGCCAC	1538
Db	1201	GATATATTATGACACGATCTACACCGGCTGTGACATGAGACCGGCTGACCCCTGCGCCAC	1260
QY	1539	GACTTGCGCGCGCTCGCAGATCGTGCCCGGGGCGCGCTACGCCATAGCGGCGCTTTCTG	1598
Db	1261	GACTTGCGCGCGCTCGCAGATCGTGCCCGGGGCGCGCTACGCCATAGCGGCGCTTTCTG	1320
QY	1589	CAGACACAGTTTGTGTAGGACACTACGCCCAAGGCCGCTTTCACCCGCGGCGCGGCGCGGAC	1658
Db	1321	CAGACACAGTTTGTGTAGGACACTACGCCCAAGGCCGCTTTCACCCGCGGCGCGGCGCGGAC	1380
QY	1659	CCCGGGCGGGGTACGGAACCGCAGCGTGCCGACACCAACGGGCTGTGTGCGCGACGAG	1718
Db	1381	CCCGGGCGGGGTACGGAACCGCAGCGTGCCGACACCAACGGGCTGTGTGCGCGACGAG	1440
QY	1719	GCCGAGGACCCGGGGGCGCGCTCGCGCGCAACGCTGTTTGTACGCGCGGCGCAACACGG	1778
Db	1441	GCCGAGGACCCGGGGGCGCGCTCGCGCGCAACGCTGTTTGTACGCGCGGCGCAACACGG	1500
QY	1779	CTGACACTTGCGGGCTCGGCTATGACACAGGCTTCGCGGCGCAACCGGCGCACG	1838
Db	1501	CTGACACTTGCGGGCTCGGCTATGACACAGGCTTCGCGGCGCAACCGGCGCACG	1560
QY	1839	CTGCTCTCTTAAGCGGCGCGGGGCGTGAAGGGCTGTGCGCTGTGCAAGCTTGCAAGCTTGCACT	1899
Db	1561	CTGCTCTCTTAAGCGGCGCGGGGCGTGAAGGGCTGTGCGCTGTGCAAGCTTGCAAGCTTGCACT	1620
QY	1889	GGCGCGCGCGCTCGGCTACTACGCGCAACCGGTGCGGCTGGGGGCGCGCGAGTCCCGCGAG	1958
Db	1621	GGCGCGCGCGCTCGGCTACTACGCGCAACCGGTGCGGCTGGGGGCGCGCGAGTCCCGCGAG	1680
QY	1959	TACTGCGGCAACCAAGTCGGGCTGTGCTGTGCTCTGTGCGCCCAACAGCGCGCGGCTCGCT	2018
Db	1681	TACTGCGGCAACCAAGTCGGGCTGTGCTGTGCTCTGTGCGCCCAACAGCGCGCGGCTCGCT	1740

QY	2019	GCGGCGCATGCGGGGCCCAATTCCTTACTGTGGCGAGAGAGCCAGAGGGCCTTGCGCCGAG	2078
Dd	1741	GCGGCGCATGCGGGGCCCAATTCCTTACTGTGGCGAGAGAGCCAGAGGGCCTTGCGCCGAG	1800
QY	2079	GCGCTGC CGCGCTGCGCGCCGGCGCGCGCGAGAGACGAAGCGCAAAGACCTGTCCGATTCC	2138
Dd	1801	GCGCTGC CGCGCTGCGCGCCGGCGCGCGCGAGAGAGCGCAAGCCCAAGACCTGTCCGATTCC	1860
QY	2139	AGCTGGATCGAAGCGCCCTCTCTCGATCAAGTCATCGACTTCAGACGACTCGCGGGATTTC	2198
Dd	1861	AGCTGGATCGAAGCGCCCTCTCTCGATCAAGTCATCGACTTCAGACGACTCGCGGGATTTC	1920
QY	2199	GAGCAGGCGCAAGCGGAGCGGATTTGCGCCGGCGCAKCGGCCGTTGCGAGTTGCTCC	2258
Dd	1921	GAGCAGGCGCAAGCGGAGCGGATTTGCGCCGGCGCAKCGGCCGTTGCGAGTTGCTCC	1980
QY	2259	CCGCTCAAGAAGCAGAGTGCTGGCCCCACGCGGAGCTGCGAGAAAGAACTCGCCAAAGACATT	2318
Dd	1981	CCGCTCAAGAAGCAGAGTGCTGGCCCCACGCGGAGCTGCGAGAAAGAACTCGCCAAAGACATT	2040
QY	2319	AGCGGCTACTATGCGTTCTACTCGACAGCTAGGCGCGCCCTGCGCGCGCGCGCGCGCG	2378
Dd	2041	AGCGGCTACTATGCGTTCTACTCGACAGCTAGGCGCGCCCTTACCGCGCGCGCGCGCGCG	2100
QY	2379	CGGCCCGGACCCCCAGCGCAGCCAGCTCTTCCCAGCTTCGCGCTCCGCCACATCC	2438
Dd	2101	CGGCCCGGACCCCCAGCGCAGCCAGCTCTTCCCAGCTTCGCGCTCCGCCACATCC	2160
QY	2439	TCCCTTGGCGCACCCAATAATTTATTTTAACCTCGATGCGCGTCTGCGAGCGCAATTAAGTCA	2498
Dd	2161	TCCCTTGGCGCACCCAATAATTTATTTTAACCTCGATGCGCGTCTGCGAGCGCAATTAAGTCA	2220
QY	2499	GGTCTCCGAGCGGATTTTTAACTTTTTTGGACAGCAGCTCTTGCAATTAGCTCACCGAC	2558
Dd	2221	GGTCTCCGAGCGGATTTTTAACTTTTTTGGACAGCAGCTCTTGCAATTAGCTCACCGAC	2280
QY	2559	CTTCAACTTTGCTGTAAA CTTTTGGTTTCTCTACTTCTTCTTGTGTGAATTACC	2618
Dd	2281	CTTCAACTTTGCTGTAAA CTTTTGGTTTGTCTTACTTCTTCTGTGTGAATTACC	2340
QY	2619	TTCCTAACATTCGCCCTCCCCCTGCTTCTTCTTACCTCCACTCTCTCTTCTGTANTGA	2678
Dd	2341	TTCCTAAC-ATTCGCCCTCCCCCTGCTGTC-TTCTTACCTCTCACTCTCTTCTGTANTGA	2398
QY	2679	AACCTTCACCTTTAGAGAGACTGTGGGAGTCTGTACAGGACAGCATTCGACCGCGC	2738
Dd	2399	AACCTTCACCTTTAGAGAGACTGTGGGAGT-CTGTACAGGACAGCGCATTC--TCGCGC	2455
QY	2739	AAGTCTCGG-CCTCCACATTAAACAATAGATGTTGACTCTTAGAACCTGAGCCACCCAGC	2797
Dd	2456	AAGTCTCGGCGCTCCACATTAAACAATAGATGTTGACTCTTAGAACCTGAGCCACCCAGC	2515
QY	2798	GCGTCTTTCTTAATCCCGAGTGGATNGATNGATNGATNGATNGATNGATNGATNGAT	2857
Dd	2516	GCGTCTTTCTTAATCCCGAGTGGATNGATNGATNGATNGATNGATNGATNGATNGAT	2574
QY	2858	TTTATGTGAA CAAGACCTGTGAAATGATGTGCACTAGTGTAAATTTATTTGTAACGATGG	2917
Dd	2575	TT---AGTGGCAAGACCTGTGAAATGATGTGCACTAGTGTAAATTTATTTGTAACGATGG	2628
QY	2918	CTAGTTTTTATTCGTCTCAGAGCACAAAACCAAGTTCATGCTTAACTTTTTTTCCTTCC	2977
Dd	2629	C-ANGTTTTATTCGTCTCAGAGCACAAAACCAAGTTCATGCTTAACTTTTTTTCCTTCC	2685
QY	2978	TTTCTTTGCTTTCTTTCTCTCTCTCAATCTTTCTCTCTCTCTTTAATTTTCTTG	3037
Dd	2686	TTTCTTTGCTTTCTTTCTCTCTCTCTCAATCTTTCTCTCTCTCTTTAATTTTCTTG	2745
QY	3038	TGAGATTAATTCTTAAGAGGCTCTAGAAACATGAAATCTCAGATGAGAGGCTTCCCA	3097
Dd	2746	TGAGATTAATTCTTAAGAGGCTCTAGAAACATGAAATCTCAGATGAGAGGCTTCCCA	2805
QY	3098	CTTCTCCCAATCCGTTGATGAAATTAATTAATTAATGTCCTTATGCACAAAATAGCTA	3157

```

Db      2806 CTTCTCCCAATCCCTTGATGATAAATTAATTAATGAGCTTAATGACACAAATAGCTA 2865
Oy      3158 AGGAGATTCACCCCAACACCTTTAAAGG 3186
Db      2866 AGGAGATTCACCCCAACACCTTTAAAGG 2894

RESULT 3
BC052737 3992 bp mRNA linear ROD 06-JUL-2005
LOCUS     Mus musculus T-box brain gene 1, mRNA (CDNA clone MGC:64688
DEFINITION
ACCESSION BC052737
VERSION    BC052737.1 GI:30931365
KEYWORDS   MGC.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS   Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
           Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
           Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
           Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
           Diachenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
           Stalcron M., Soares M.B., Bonaldo M.F., Casavant T.L.,
           Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
           Carninci P., Frange C., Raha S.S., Loguelano N.A., Peters G.J.,
           Abramson R.D., Mullany S.J., Bosak S.A., McKwan P.J.,
           McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
           Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
           Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
           Fahey J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko Y.,
           Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
           Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D.,
           Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
           Butlerfeld Y.S., Krzywinski M.I., Skalska U., Smalins D.E.,
           Scherck A., Schein J.E., Jones S.J. and Marra M.A.
           Mammalian Gene Collection Program Team
           Generation and initial analysis of more than 15,000 full-length
           human and mouse cDNA sequences
           Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
           12477932
           2 (bases 1 to 3992)

JOURNAL   NIH MGC Project
PUBMED    Direct Submission
CONSRTM   Submitted (19-MAY-2003) National Institutes of Health, Mammalian
TITLE     Gene Collection (MGC), Bethesda, MD 20892-2590, USA
JOURNAL   NIH-MGC Project URL: http://mgc.nhl.nih.gov
REMARK    Contact: MGC help desk
COMMENT   Email: cgabs-remail.nih.gov
           Tissue Procurement: Dr. Jim Lin, University of Iowa
           CDNA Library Preparation: M. Bento Soares, University of Iowa
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
           DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
           Thomas U. Casavant.
           Web site: http://genome.uiowa.edu
           Contact: Bento-soares@uiowa.edu; tom-casavant@uiowa.edu
           Bonaldo M.F., Akabogu I., Bait T., Bait J., Crouch K., Davis A.,
           Fisher K., Keppel C., Kucaba T., Lebeck M., Melo A., Schaefer K.,
           Schetz T., Smith C., Snir E., Tack D., Trout K., Walters J.,
           Casavant T., Soares M.B.
           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
           Series: Plate: Row: Column: 0
           This clone was selected for full length sequencing because it
           passed the following selection criteria: matched mRNA gi: 3438150.
           Location/Qualifiers
           1. 3992

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/secretin="CS7BL/6"
/db_xref="taxon:10090"
/clone="MGC:64688 IMAGE:6825394"
/tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
/clone_id="NIH BMAP_FY0"
/lab_host="DH10B"
/notes="vector: pYX-ASC"
1. 3992
/gene="Tbr1"
/db_xref="GeneID:21375"
/db_xref="MGI:107404"
256 2301
/gene="Tbr1"
/codon_start=1
/product="T-box brain gene 1"
/protein_id="AAH52737.1"
/db_xref="GI:30931366"
/db_xref="GeneID:21375"
/db_xref="MGI:107404"
/translation="MQLHECLSPSILMSKRFVNVSSYPHSGSEVLAHDPITSTD
NLEKSSPLKRTKRTMTQSDTDFPDSKDSFGDVORSLSEVLDVSELRHSFPGSAA
DRYLSQSSQPSQATAPASMFYPYSHQEPAPAFSLGSPERYAAHHPVITNGANSL
LNSSPQSPAGYPYPOQGYHGYOGAPFYQSTQGLVPGKQVYLCNPLDKFL
RHQEMETITKQGRMMPPLSPFNISGLDPLTAHYNLFVNYLADPNHMFQGGKVPCKG
ADTVQGNRYVNHDSRNTGAMHROBISRGKLTNNKSNNGNMVYQSLHKYQ
PRLHVEVNEDEGTDTSOPGVQVPTTPETPTPIATVATYQNTDIQKIDNPAPKGR
DNYDTITVGCMDRLTPSPDSPPSQVPGARVMAASFLDDQVSNYAKRFHPGAG
AGPBGDPSVPHNGLSPQASDPGAPQAPQVAPANNRDLFASVATATDIFAG
NATULLSAGVAVALPLQAGCTGRPLGYADSPGWCARSPQYCGAKSGSVLPCW
NSAAABRMAGANVLEBEAGLEAERSPLAPALDEKAPKQLSNSETSPSISKID
SSDSGLYEQARRRISPDITVSSSSSLKSEVLAQNDCKNCKADIGYGYGFVSHS"

ORIGIN
Query Match 78.2%; Score 2490; DB 9; Length 3992;
Beet Local Similarity 90.0%; Pred. No. 2,8e-311;
Matches 2870; Conservative 0; Mismatches 245; Indels 73; Gaps 16;

Oy      47 CAGCTAACCAATGAGAGAGCCGTTTCATCGCGTGACACAGACGAAATCAAAAGTAG 106
Db      1 CAGCTAACCAATGAGAGAGCCGTTTCATCGCGTGACACAGACGAAATCAAAAGTAG 60

Oy      107 CTTTTCGATTGTCGCAATAGTCAATTGGCAATCTTCTCCAGGAAAAAAA 166
Db      61 CTTTTCGATTGTCGCAATAGTCAATTGGCAATCTTCTCCAGGAAAAAAA 120

Oy      167 GTAATCAACCTTTGAGAGAGATTTGCTGGTGAAGCTTTCTGTAGTAGAGGGGT 226
Db      121 GTAATCAACCTTTGAGAGAGATTTGCTGGTGAAGCTTTCTGTAGTAGA-GGGGT 179

Oy      227 CTGTGATTTCTAGTTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGGAGTG 286
Db      180 CTGTGATTTCTAGTTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGGAGCGG 239

Oy      287 TTCAGGTTCTAGAGCTAGACACTGAGAGACCTGCTTCTCTTATCATGCTGCCAA 346
Db      240 TTCAGGTTCTAGAGCTAGACACTGAGAGACTGCTTCTCTTATCATGCTGCCAA 299

Oy      347 GAAATTTCTCAATGTGAGAGAGAGCTACCAATTCAGGCGAGTCGAGCTTGTCTTGA 406
Db      300 GAAATTTCTCAATGTGAGAGAGAGCTACCAATTCAGGCGAGTCGAGCTTGTCTTGA 359

Oy      407 CGATCATCCCATATCTGAGACCTGAGCAACCTGAGAGAGACTTCACTTTAAAAAAAT 466
Db      360 TGATCATCCCATATCTGAGACCTGAGCAACCTGAGAGAGACTTCACTTTAAAAAAAT 419

Oy      467 TACAGGGGAGAGAGAAATCACTAGATACAGAAATTTCTGACTCAAGAGACTCAC 526
Db      420 TACAGGGGAGAGAGAAATCACTAGATACAGAAATTTCTGACTCAAGAGACTCAC 479

Oy      527 AGGGAGCTCCAGAGAAATCACTCTCTGCTCTTGGACGGGGTCTGAGACTTCGTCA 586

```

Db 480 AGGGAGCGTCCAGAGAAAGTAACTCTCTCTGTGAGCGGGGTCCTGAGCTTCGCA 539
Qy 587 CAGTTTCATGAGCTCTGCTGAGATGCTACTCTCTCTGATGTCAGCCAGCCAGATC 646
Db 540 CAGTTTCATGAGCTCTGCTGAGATGCTTAACTTACTCTCTCTGATGTCAGCCAGCCAGATC 599
Qy 647 TGGCGGCACTGCTCCAGTGGCCATGTTGCCGTACCCCGGCGAGCAAGGAGCGGCGACCC 706
Db 600 TGGGCAACCGCTCCAGTGGCCATGTTGCCGTACCCCGGCGAGCAAGGAGCGGCGATCC 659
Qy 707 GCGCTTCTCATCGGAGCGCCCTAGCCGCTACATGAGCCCAACCCCGGTCTATCACCAGG 766
Db 660 GCGCTTCTCATCGGAGCGCCCAAGTCCGCTACATGAGCCCAACCCCGGTCTATCACCAGG 719
Qy 767 AGCTTACAACAGCT 826
Db 720 AGCTTACAACAGCT 779
Qy 827 CTACCCACAGAGTACGAGGCACTCTCTACCAAGAGAGCTCGTTCTACAGATTCTCTCTCAC 886
Db 780 CTACCCACAGAGTACGAGGCACTCTCTACCAAGAGAGCTCTCTCTCTCTCTCTCTCTCAC 839
Qy 887 CCAGCCGAGGCTGTGTGCGCGGCAAGCAAGGTGTACTGTGTGCAACAGGCCCTTTGAGCT 946
Db 840 CCAGCCGAGGCTGTGTGCGCGGCAAGCAAGGTGTACTGTGTGCAACAGGCCCTTTGAGCT 899
Qy 947 GAAATTTTACCGGCAACCAACGGAATGATCATACCAACAGGGAGCGGCTTTCC 1006
Db 900 GAAATTTTACCGGCAACCAACGGAATGATCATCTAAACAGGGAGCGGCTTTCC 959
Qy 1007 TTTTAAATTTTAACTTTCTGATCTGATCCAGCGCTCATTAATTTTGTGGA 1066
Db 960 CTTTGTGATTTTAACTTTCTGATCTGATCCAGCGCTCATTAATTTTGTGGA 1019
Qy 1067 TGTGATTTTGTGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGAGTTCTTGTGCG 1126
Db 1020 TGTGATTTTGTGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGAGTTCTTGTGCG 1079
Qy 1127 CAAAGCGAGCACTATGTCAGAGAAATCGGCTTATATGCAATCCGAGTTCCCGCAACAC 1186
Db 1080 CAAAGCGAGCACTATGTCAGAGAAACCGGCTTATATGCAATCCGAGTTCCCGCAACAC 1139
Qy 1187 TGGGGCTCATGAGTGGCGCAAGAAATCTTTTGGAAAAATTTAACTTAACTTAACTTAA 1246
Db 1140 TGGGGCTCATGAGTGGCGCAAGAAATCTTTTGGAAAAATTTAACTTAACTTAACTTAA 1199
Qy 1247 AGGAGCTTCAAAATAAATAATGAGTGGTGTGTTTAACTGCTTGTGCAAAATGACAGCC 1306
Db 1200 AGGAGCTTCAAAATAAATAATGAGTGGTGTGTTTAACTGCTTGTGCAAAATGACAGCC 1259
Qy 1307 CCGCTGCTGATGTTGGAAGTGAAGAGACGCGCAAGAGCACTTACCGAGCCCGGCG 1366
Db 1260 CCGCTGCTGATGTTGGAAGTGAAGAGAGTGGCAAGAGCACTTACCGAGCCCGGCG 1319
Qy 1367 CGTGGAGAGTTCATCTTCCCTGAGACTGATTCATCCGCGCTACCGGCTTACAGAAC 1426
Db 1320 AGTCCAGAGTTTCACTTTCCGAGACTGATTCATCCGCTGCTACCGGCTTACAGAAC 1379
Qy 1427 GGATATTACAACTGAAAAATAGATCACAACTTTTGTGAAAAAGATTTCCGGGATTAATTA 1486
Db 1380 GGATATTACAACTGAAAAATAGATCATTAACCTTTGTGAAAAAGATTTCCGAGATTAATTA 1439
Qy 1487 TGACACGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAAAGATCGCC 1546
Db 1440 TGACACGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAAAGATCTTCC 1499
Qy 1547 GGGCTGCGAGATCGTGGCGCGGGGCGCGCTAGCGCAATGGCGGGCTTTTCTGTCAGAGCA 1606
Db 1500 GGGCTGCGAGATCGTGGCGCGGGGCGCGCTAGCGCAATGGCGGGCTTTTCTGTCAGAGCA 1559
Qy 1607 GTTCTGAGCAACTACGCGCAAGCGCGCTTCCACCGGCGCGGCGCGCGCGCGCGCG 1666
Db 1560 GTTCTGAGCAACTACGCGCAAGCGCGCGCTTCCACCGGCGCGCGCGCGCGCGCGCGCG 1619

Qy 1667 GGGTACGGAACCGAGCGTGGCGGACACAAAGGGCTGTGTGCGCGGAGAGGCGGAGGA 1726
Db 1620 GGGGACGGACCGGAGCGTGGCGGACACAAAGGGCTGTGTGCGCGGAGAGGCGGAGGA 1679
Qy 1727 CCGGGGCGCGCGCTGCGCGGCAACGCTGTTTGTGAAGCGCGGCGCAACACCGGCTGAGCTT 1786
Db 1680 CCGGGGCGCGCGCTGCGCGGCAACGCTGTTTGTGAAGCGCGGCGCAACACCGGCTGAGCTT 1739
Qy 1787 CGCGGCTTGGGCTTATGACACGCGCAAGGACTTGGCGGGCAAGCGGGCCAGCTGCTTC 1846
Db 1740 CGCGGCTTGGGCTTATGACACGCGCAAGGACTTGGCGGGCAAGCGGGCCAGCTGCTTC 1799
Qy 1847 TTAAGCGGCGGAGGCTGAAGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1906
Db 1800 GTATGCGGCGGAGGCTGAAGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1859
Qy 1907 GCTGAGCTACTAGCGCGACCGCTGAGGCTGAGGCGCGGAGGCTGAGGCTGAGGCTGAGG 1966
Db 1860 GCTGAGCTACTAGCGCGACCGCTTGGGCTGAGGCGCGGAGGCTGAGGCTGAGGCTGAGG 1919
Qy 1967 CACCAATCGGAGCTCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 2026
Db 1920 CGCCAGTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 1979
Qy 2027 GCGCGGCGCAATCCCTACCTGGGCGAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGG 2086
Db 1980 GCGCGGCGCAATCCCTATCTGGGCGAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGG 2039
Qy 2087 GCTGCGCGCGGCGCGCGCGGAGGCGGCAAGGCGGCAAGGCTGTCCGATTTCAAGTGGAT 2146
Db 2040 GCTGCGCGCGCGCGCGCGCGGAGGCGGCAAGGCGGCAAGGCTGTCCGATTTCAAGTGGAT 2096
Qy 2147 GGAAGCGCTCTCTGATCAAGTTCATGACTGAGGAGTCTGGGGATTTACGAGCGG 2206
Db 2097 GGAAGCGCTCTCTGATCAAGTTCATGACTGAGGAGTCTGGGGATTTACGAGCGG 2156
Qy 2207 CAAAGCGAGGCGGATCTGCGCGGCGAGCGCGGCTGCGAGGTTGCTCCCGGCTCAA 2266
Db 2157 CAAAGCGAGGCGGATCTGCGCGGCGAGCGCGGCTGCGAGGTTGCTCCCGGCTCAA 2216
Qy 2267 GAGCGAGTGTGCGCGGCGGAGCTGCGAGAAAGAACTGCGCGCAAGGACATTAAGCGGCTA 2326
Db 2217 GAGCGAGTGTGCGCGGCGGAGCTGCGAGAAAGAACTGCGCGCAAGGACATTAAGCGGCTA 2276
Qy 2227 CTATGCTTCTACTGCGACAGCTAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 2386
Db 2277 CTATGCTTCTACTGCGACAGCTAGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 2336
Qy 2387 ACCCGAGCGAGCGCGCTGACAGCTTTCGCGAGCTGCGCGCTGCGCGCAACTCTCTTGGC 2446
Db 2337 GCG 2384
Qy 2447 CAGCCACTCATTTTATTTGACCTCTGATGAGCGGCTGTGACAGATTAAGTCAAGGCTTCCG 2506
Db 2285 CAGCTGCGCTCTTACATTTGACCTCTGATGAGCGGCTGTGCGCGGAGTAAAGTCAAGGCTTCTC 2442
Qy 2507 AGCGTGAATTTTAACTTTT-----TGACAGCACTCTCTG 2542
Db 2443 ACTATGATTTTAAAACTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTG 2502
Qy 2543 CAATTTACTACCGAGCTTCAACTTTGCTGTAACCTTTTGTGTTTCTTTCTTTCTTTCTTTCT 2602
Db 2503 CAGTTAGGCAACCGAGCTTGAACCTGCTGTAACCTTTGAGTTTTCATTAATT-----TTC 2558
Qy 2603 TTTCTGAGGATTAATCTCTTAC-----AATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2656
Db 2559 GCTCTGAGGATTAATTAATCT 2618
Qy 2657 CTACTTCTTTCTTTGTAATGAATCTTTTCACTTTTGAAGACTTGGGCACTGCTGTGAG 2716
Db 2619 CTTCTCTTTCTTTGTAATGAATCTTTTCACTTTTGAAGACTTGGGCAATCTGTGAG 2678

Db	1	CTCTTCTCCACGGGAAAAAAAGTAATCAAACTTTTGAGAAAGCATTTCTGTTGGA	60
OY	203	GTGCTTTCGTGCTAGTGAAGGGGTCGTGGAATTTCTAGTTATGATAAATAGACTTTAA	262
Db	61	GTGCTTTCGTGCTAGTGA - GGGGTCGTGGAATTTCTAGTTATGATAAAGGACTTTAA	119
OY	263	AAACCAAGGAGCGGAGGGCGAGTGTTCAGGTTCTAGACTATGCAAGCTGAGCACTGCCT	322
Db	120	AAACCAAGGAGCGGAGGGCGAGCGTTCAAGTCTTAGACTATGCAAGCTGAGCACTGCCT	179
OY	323	TTTCTCCTTATCATGCTCTCCAAAGAAATTTCTCAATGAGAGCACTACCAATTC	382
Db	180	CTTCTCTTCTATCATGCTCTCCAAAGAAATTTCTCAATGAGAGCACTACCAATTC	239
OY	383	AGCGCGATCCGAGCTTGTCTTGCAAGATATCCATTAATCTGACACATGCAAACTTGA	442
Db	240	GGGCGGATCTAGCTTGTCTTGCAAGATATCCATTAATCTGACACATGCAAACTTGA	299
OY	443	GAGAAATTACCTTTGAAAAAATTAACAAGGGGATGACGATCATGTCAGATACAGACA	502
Db	300	GAGAAATTACCTTTGAAAAAATTAACAAGGGGATGACGATCATGTCAGATACAGACA	359
OY	503	TTTTTCCTATCCCAAGGACTCAACAGGGGAGTCCAGAGAAAGTAACCTCTCCGCTT	562
Db	360	TTTTTCCTATCCCAAGGACTCAACAGGGGAGTCCAGAGAAAGTAACCTCTCTGCTTT	419
OY	563	GGACGGGGTCTCTGAGCTTCGTCAAGTTTCGATGGCTCTGTCGATATGCTACTTCT	622
Db	420	GGACGGGGTCTCTGAGCTTCGTCAAGTTTCGATGGCTCTGTCGATATGCTACTTCT	479
OY	623	CTCTCAGTCCAGCCAGCCACAGTCTGCGGCCACTGCTCCAGTGCATGTTCCGTAACC	682
Db	480	CTCTCAGTCCAGCCAGCCACAGTCTGCGGCCACCGCTCCAGTGCATGTTCCGTAACC	539
OY	683	CGGCGAGGACGGAGCGGACCCCGCTTCCTCAATGGGCAAGCCCTACCGCTTCAATGGC	742
Db	540	CAGCCAGGACGGAGCGGACCGGCTTCCTTCCTCAATGGGCAAGCCCTACCGCTTCAATGGC	599
OY	743	CCACCACCGGCTATCACCAAGGAGCTTCAACAGCTCTGTCGCACTCTCTGCGCGCA	802
Db	600	CCACCACCGGCTATCACCAAGGAGCTTCAACAGCTCTGTCGCACTCTCTGCGCGCA	659
OY	803	GGGATACCCACGCGCGGCTACCCCTACCCACAGGATACGCGCACCTCTTACCAAGAGC	862
Db	660	GGGATACCCACGCGCGGCTACCCCTACCCACAGGATACGCGCACCTCTTACCAAGAGC	719
OY	863	TCCGTTTACAGTCTCTCCACCCAGCCCGGGGCTGATGCGCCGCAAGCAAGCATGCTA	922
Db	720	CCCTTTTACAGTCTCTCCACCCAGCCCGGGGTTGATGCGCCGCAAGGAGCATGTA	779
OY	923	CCTGTGCAACAGGCGCCCTTGTGGCTGAATTTTACACGGGACCAACAGAGATGATCATAC	982
Db	780	CCTGTGTCAACAGGCGCCATTTGTGGCTGAATTTTCAATGGGATCAACAGAGATGATCATAC	839
OY	983	CAAAAGGAGAGCGCATGTTCTTTTAAATTTTAAATTTCTGATCTGATCCAC	1042
Db	840	TAAAAGGAGAGCGCATGTTCTTTTAAATTTTAAATTTCTGATCTGATCCAC	899
OY	1043	GGCTATTACAATTTTTTGTGAGATGATTTTGGCGGATCCCAATCATGGAAGTTTCA	1102
Db	900	CGCTATTACAATTTTTTGTGAGATGATTTTGGCGGATCCCAATCATGGAAGTTTCA	959
OY	1103	AGGAGGCAATATGGGTTCTGTGGGCAAAAGCGGACCAATGTGCAAGGAAATCGGGTCTA	1162
Db	960	AGGAGGCAATATGGGTTCTGTGGGCAAAAGCGGACCAATGTGCAAGGAAATCGGGTCTA	1019
OY	1163	TATGATCCGAGTTTCCCAACACTGCGGCTACTGATATGCGCCAAAGAAATCTTTTGG	1222
Db	1020	TATGATCCGAGTTTCCCAACACTGCGGCTACTGATATGCGGCAAGAAATCTTTTGG	1079
OY	1223	AAATTTAAACTTACGAACCAACAAAGAGCTTCAATTAACAATGGGCAAGATGATGTTT	1282
Db	1080	AAATTTAAACTTACGAACCAACAAAGAGCATCAACAAACAATGGGCAAGATGATGTTT	1139

QY	1283	ACAGTCTTGCACAAGTACCAGCCCCCGCTGCATGTGTGGAAAGTGAACGAGACGGCAC	1342
Db	1140	ACAGTCCCTGCACAAGTACCAGCCCCCGCTGTGCAGTGTGTGAAGTGAATAGAGTGTGCAC	1199
QY	1343	GGAGGACACTAGCCAGCCCGGCGCGTGCAGACGGTTACATTTCTCTAGACTCAGTTTCAT	1402
Db	1200	AGAGGACACCAGCCAGCCAGGCGCAGTGCAGACGTTCACTTTTCCGGAGACTCAGTTTCAT	1259
QY	1403	CGCGGTCAACGCGCTACCAAGAACACGAGTATTACACAACCTGAATAATAGATCAACCCCTT	1462
Db	1260	CGGTGTCAACCGCTACCAAGAACACAGGATATTACACAACCTGAATAATAGATCAACCCCTT	1319
QY	1463	TGCAAAAGGATTTGGGGATATTATTGACAGACTCTACACCGGCGGTGACATGTGACCGCT	1522
Db	1320	TGCAAAAGGATTTGAGGATTAATATGACAGATCTTACACGGGCTTGCAATGTGACCGCT	1379
QY	1523	GACCCCTCTCGCCCAACGACTCGCCGCGCTCGCAGATCGTGCCCGGAGGCCGCTACCGCAT	1582
Db	1380	GACCCCTGTGCGCCCAACGACTCTCGCGCGCTCGCAGATCGTGCCCGGAGGCCGCTACCGCAT	1439
QY	1583	GGCGGACTCTTTCTGTGACAGACCAAGTTCTGTAGGAACTACAGCCCAAGGCCCGCTTCCACCC	1642
Db	1440	GGCGGACTCTTTCTGTGACAGACCAAGTTCTGTAGGAACTACAGCCCAAGGCCCGCTTCCACCC	1499
QY	1643	GGGCGGAGCGCGGAGCCCGCGGCGGAGGTACGGAACGAGCGTGCCGACACCAACGAGACT	1702
Db	1500	GGGCGCGCGCGCGGACTCCCGGAGCGGAGCGAGAACCGAGAGGTGCGCACACCAACGAGACT	1559
QY	1703	GCTGTGCGCGCAGACAGGCGCCAGGACCCGGGCGCGCCCTCGCCGACAGGCTGTGTGTGAC	1762
Db	1560	GCTGTGCGCGCAGACAGGCGCCAGGAGACCCGGGCGCGCGCTCGCCGACAGGCTGTGTGTGAC	1619
QY	1763	GCGGCGCCACAACCGGCGTGGACTTCGCGGCGCTCGGCTATTGACACCGGCAACGAGACTTCGC	1822
Db	1620	GCGGCGCCACAACCGGCTGTGACTTCGCGGCGCTCGGCTATTGACACCGGCAACGAGACTTCGC	1679
QY	1823	GAGGACAACGCGGCAACGCTGTCTTTCACGCGGCGGCGGCGGTGAAGGCGCTGCCGCTGCA	1882
Db	1680	GCGGACAACGCGGCAACGCTGTCTGTGTGTGTGCGGCGGCGGCGGTGAAGGCGCTGCCGCTGCA	1739
QY	1883	GAGCTGCAGAGCTGTGACTGGGCGGCGGCTGCGCTACTACGCGGACCCGCTCGGCGTGGGCGG	1942
Db	1740	GAGCGCGGAGCTGTGACAGGCGGCGGCGCTGCGCTACTACGCGGACCCCTTCGCGGCTGGGCGG	1799
QY	1943	CGGCAGTGCCCCCGGAGTACTGCGGCGCACCAAGTCGAGGCTCGGCTGTGCGCTGTGCGCCAA	2002
Db	1800	GCGGAGCGCCCCCGGAGTACTGCGGCGCACAGTCGAGGCTCGGCTGTGCGCTGTGCGCCAA	1859
QY	2003	CAGCGCGCGCGGCGCGCGCGCGCATGCGCGCGCATATCCTACTGGGCGGAGAGGCGCGA	2062
Db	1860	CAGCGCGCGCGGCGCGCGCGCGCATGCGCGCGCATATCCTACTGGGCGGAGAGGCGCGA	1919
QY	2063	GGGCGCTGCGCGCGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGAGGACGCGCAAGCCCAA	2122
Db	1920	GGGCGCTGCGCGCGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGAGGACGCGCAAG	1976
QY	2123	GGAACCTGTCCGATTCAGCTGGATCGAGACGCGCTCTCGATCAAGTCCATCGACTCCAG	2182
Db	1977	GGAACCTGTCCGATTCAGCTGGATCGAGACGCGCTCTCTCGATCAAAATCCATCGACTCCAG	2036
QY	2183	GCACTCGGAGGATTTTACAGACAGGCGCAAGCGGAGGCGGATCTGCGCGGCGTGCACAGCCGCT	2242
Db	2037	GCACTCGGAGGATTTTACAGACAGGCGCAAGCGGAGGCGGATCTGCGCGGCGTGCACAGCCGCT	2096
QY	2243	GTCCGAGAGTTCGTCCCGCTCAAGAGCGAGGCTGCGCCACGCGGAGCTTCGAGAAAGAA	2302
Db	2097	GTCTGAGAGCTGTGTCGCCGCTCAAGAGGAGAGGTCTGCGCCACGCGGAGCTTCGAGAAAGAA	2156
QY	2303	CTGCGCCCAAGACATTAGCGGCTACTATGCTTTTACTCGACACGCTAAGCGCGCCCTGC	2352
Db	2157	CTGCGCCCAAGACATTAGCGGCTACTATGCTTTTACTCGACACGCTAAGCGCGCGCACGCC	2216

Db	300	CCTGACTCGAAGAACTCACACAGGGAGCGTCCAGAGAACTAACTCTCTCTGTCTTGAGC	359
Oy	567	GGGGCTCTGAGCTTCGTCAAGATTTCGATGAGTCTGTGTCAGATTCGCTACTCTCTCT	626
Db	360	GGGGCTCTGAGCTTCGTCAAGATTTCGATGAGTCTGTGTCAGATTCGCTACTCTCTCT	419
Oy	627	CAGTTCAGCCACGACCAAGTCTGCGGCCACTGTCTCCAGTGGCAGTTTCCGTACCCGGC	686
Db	420	CAGTTCAGCCACGACCAAGTCTGCGGCCACCGCTCCAGTGGCAGTTTCCGTACCCGAC	479
Oy	687	CAGCAGCGAACCAGGCGCACCCCGCCTTCACATCGGCGACCCCTAGCCGCTACATGAGCCAC	746
Db	480	CAGCAGCGAACCAGGCGCATTCGCGCTTCCTCATTGCGAGCCCAAGTCTGCTAATGACCCAC	539
Oy	747	CACCGGCTCATCACCAAGAGAGCCCTACAAAGCCCTCTGTCCAACTCTCTCGCCAGAGGA	806
Db	540	CACCGGCTCATTCACCAAGAGAGCTTACAAGACCTGTGTCCAACTCTTGGCCGCAAGGAC	599
Oy	807	TACCCCAAGCGCGCGCTACCCCTACCCACAGCAGTACGCGCACTCTTACCAAGAGCTCG	866
Db	600	TACCCCAAGCGCGCGCTACCCCTACCCACAGCAGTACGCGCACTCTTACCAAGAGCCCT	659
Oy	867	TTCTACCAAGTTCCTCTCCACCCAGCCCGAGTGTGTGCCCGAAGGCCCAAGTAACTGT	926
Db	660	TTCTACCAAGTTCCTCTCCACCCAGCCCGAGTGTGTGCCCGAAGGCCCAAGTAACTGT	719
Oy	927	TGCAACAGAGCCCTTGTGGCTGAATTTACCGGACACGAAAGAGATGATCATCACAA	986
Db	720	TGCAACAGAGCCACTTGTGGCTGAATTTATCTGGCATTCAAAGAGATGATCATCATTA	779
Oy	987	CAGGAGAGCGCATGTTTCTTTTAAAGTTTAAATTTCTGTCTCGATCCACGCGCT	1046
Db	780	CAGGAGAGCGCATGTTTCTTTTAAAGTTTAAATTTCTGTGTCTCGATCCACGCGCT	839
Oy	1047	CATTACATATTTTGTGATGTGATTTTGTGGGATCCCATTCCTGAGGTTTCAAGGA	1106
Db	840	CATTACATATTTTGTGATGTGATTTTGTGGGATCCCATTCCTGAGGTTTCAAGGA	899
Oy	1107	GGCAATGAGGTTCTTGTGGGCAAAAGCGACACCATGTGCAAGAAATCGGGTCTATATG	1166
Db	900	GGCAATGAGGTTCTTGTGTGCAAAAGCGACACCATGTGCAAGAAACGGGTCTATATG	959
Oy	1167	CATCCGATTCGCCCAACACTGAGGCGCTCATGTGATGCGCCAAAGAAATCTTTTGGAAA	1226
Db	960	CATCCGATTCGCCCAACACTGAGGCGCTCATGTGATGCGGCAAGAAATCTTTTGGAAA	1019
Oy	1227	TTTAAATCTTACGAACAACAAGAGACTTCAATATACATGGCGAGATGCTGTTTACG	1286
Db	1020	TTTAAATCTTACCAACAAAGGAGCATTCAAACAATGGGCAATGTGTGTTTACG	1079
Oy	1287	TCTTTGCAACAATACAGGCCCGCGCTGATGTGTGGAAGTGAAGAGAGCGGACGAG	1346
Db	1080	TCTTTGCAACAATACAGGCCCGCGCTGATGTGTGGAAGTGAAGAGAGTGGCACAGG	1139
Oy	1347	GACACTAGCAGCCCGCGCGCTGACAGAGTTCACTTTCCCTGAGACTCACTTCAATGCGC	1406
Db	1140	GACACTAGCAGCCCGCGCGAGTCCAGAGCTTCACTTTCCGAGAGCTCACTTCAATGCGC	1199
Oy	1407	GTCACCGCTTACCAAGAACGAGATTTACACAACCTGAATATGATCAACAACCTTTTGA	1466
Db	1200	GTCACCGCTTACCAAGAACGAGATTTTACACAACCTGAATATGATCAACAACCTTTTGA	1259
Oy	1467	AAAGATTTCCGGGATTAATTATACAGATCTACACCGGTGTGACATGAGACGCGCTGAC	1526
Db	1260	AAAGATTTCCGGGATTAATTATACAGATCTACACCGGTGTGACATGAGACGCGCTGAC	1319
Oy	1527	CCCTGCGCCACGACTCGCGCGCTCGCAGATCTGTGCCCGGCGCGCTACGCGCAATGCGC	1586
Db	1320	CCGTGCGCCACGACTCTTCGCGGCTCGCAGATCTGTGCCCGGCGCGCTACGCGCAATGCGC	1379
Oy	1587	GGCTCTTTCCTGACGAGCAATTCGTGAGCACTACCGCAAGGCCCGCTTCCACTCGGCGC	1646
Db	1380	GGCTCTTTCCTGACGAGCAATTCGTGAGCACTACCGCAAGGCCCGCTTCCACTCGGCGC	1439

Oy	1647	GCGGGCCCCGGGCCCCGGGCGGTACGAGACCCGACAGCTGCTCCGACACACAAACGGGCTGCTG	1706
Db	1440	GCCGGCCCGGGATCCCCGGGCGGGGCGACCGAGCCGAGCTGGCCGACACAAACGGGCTGCTG	1499
Oy	1707	TCCGCGGAGACAGGCGCGAGAGACCCGGGCGCCGCTCCGCGCAACGCTGGTTGTGACGCGC	1766
Db	1500	TCCCCGACAGAGCGCGAGAGACCCGGGCGCCCGCTCCGCGGAGCGTGGTTGTCAAGCGC	1559
Oy	1767	GCCAAACAACCGGCTGACCTTCGGGCGCTCGGCTTATGACACGGGCAACGACCTTCGGGGG	1826
Db	1560	GCCAAACAACCGGCTGACCTTCGGGCGCTCCGCGCTACGACACGGGCAACGACCTTCGGGG	1619
Oy	1827	AACGGGGCCACGCTGCTCTTTACGGGGCGGGCGGTGAAGGGGCGTGGCGCTGACAGCT	1886
Db	1620	AACGGGGCCACGCTGCTGTGTGTATGGCGCGGGGCTGAAGGGGCTGGCTTGGAGGCC	1679
Oy	1887	GCAAGCTGCACTGGCGCGCCGCTCGGCTACTACGCGGACCGTCGCGGCTGGGGCGCCG	1946
Db	1680	GCGGGCTGCAAGGGCGCGCCGCTGGCTACTACGCGGACCTTCGGGCTGGGGCGCGG	1739
Oy	1947	AGTCCCCCGAGTACTGGCGCACCAATCGGGCTCGGTGTGCTGCCCTGCTGGGCCAAGC	2006
Db	1740	AGCCCCCGAGTACTGGCGCGCCAAATCGGGCTCGGTGTGCCCTGCTGTCGCCAAGC	1799
Oy	2007	GCCGCGGCGCGCGGCGGCATGCGCGGGCGCAATCCCTACCTGGGCGAGAGGCGGAGG	2066
Db	1800	GCCCGCGCGCGCGGCGGCATGCGCGGGCGCAATCCCTATCTGGGCGAGAGGCGGAGG	1859
Oy	2067	CTGGCGCGCGGCGCTCGCGCTGCGCCCGGCGCGCGCGCGGCGGACCGCAAGCCCAAGAC	2126
Db	1860	CTGGCGCGCGGCGCTCGCGCTGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1916
Oy	2127	CTGTCCGATTCACACTGGATGAGACGCGCCCTCGATTCAGATCATCGACTCGAGGAC	2186
Db	1917	CTGTCCGATTCACACTGGATGAGACGCGCTCTCTCATCAATTCATCGACTCGAGGAC	1976
Oy	2187	TCCGGGATTTTACGAGACGGCCAAAGCGGAGCGGATCTCGCGCGCGACGCGCCGTGTCC	2246
Db	1977	TCCGGGATTTTACGAGACGGCCAAAGCGGAGCGGATCTCGCGCGCGGTGACAGCGCGGTCT	2036
Oy	2247	GAGAGTTCGTCGCCGCTCAAGAGGAGGATGCTGGCCGAGGGGACCTCGAGAGAACTGC	2306
Db	2037	GAGAGCTGTGTCGCCGCTCAAGAGGAGGATGTCGCCGAGGGGACTCGAGAGAACTGC	2096
Oy	2307	GCCAAAGACATTAGCGGCTACTATCGGCTTCTACTCGACACGCTAGGCGGCGCCCTCGCG	2366
Db	2097	GCCAAAGACATTAGCGGCTACTATGGCTTCTACTCGACACGCTAGGCGGCGGCTCG	2156
Oy	2367	CGGGCGCGCGCGCGCGGACCCCGACCGCCCTCAAGCTCTTCGCCGACTCGGC	2426
Db	2157	GCCGCGCGCGCGCGCGGACCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2204
Oy	2427	TCCGACACCTGCTCTTGGCGGACCCGACCTATTTATTTGACCCCTGATGGCGCTCGAG	2486
Db	2205	TTCGAGCGCGCTTCTAGCGGACTCGGCTCTTCACTTGAACCTCTGATGACCGTCTCGG	2263
Oy	2487	GGAATTAAGTCAGGCTCTCCGAGCGGTGATTTTAACTTTT-----	2526
Db	2264	GG-ATAAGTCAGGCTCTCTCACTATGATTTTAAACTCTTTCTTTTCTTTCTTTCTTTTC	2322
Oy	2527	----TGACAGGACATCTGCAATTACCTACCGGACCTTCAACTTGTCTGTAAACCTTTT	2582
Db	2323	TCTCTACACAGCCTTCTGTGAGTTAGGACCGGACCTTGAACCTGTGCTGTAAACCTTGT	2382
Oy	2583	GATTTCCTACTTACTCTTCTTCTGTGAGATTACTCTTCAACAATC-----CCCTCC	2636
Db	2383	GATTTCCTACTT-----TTGCTGTGAGGTTATGATCTCTCCGTCTTTTTCACACCC	2438
Oy	2637	CCTGCTCTTCTCTTACCTCTTACTTCTCTTTCTTGTAAAGAACTCTTACACTTTTAGA	2696
Db	2439	CTTCTCTCTTGGCCCACTCATCTCTCTTCTCTTGTAAATGAACCTTCAACTTTTAGA	2498

[illegible]

OY	345	AAAGAAATTTTCATATGTGAGAGAGAGGTACCAATTCAGGGCGATCCAGCTGTCTTG	404
Db	121	AAAGAAATTTTCATATGTGAGAGAGGTACCAATTCAGGGCGATCCAGCTGTCTTG	180
OY	405	CACGATCATCCATATATCTCGACCACTGAGAGAAATTCACTTTGAAAAA	464
Db	181	CACGATCATCCATATATCTCGACCACTGAGAGAAATTCACTTTGAAAAA	240
OY	465	ATTACAGAGGGGATGACGAATCACTGAGATACAAATTTTCTGACTCCAAAGACTCA	524
Db	241	ATTACAGAGGGGATGACGAATCACTGAGATACAAATTTTCTGACTCCAAAGACTCA	300
OY	525	CCAGGGGACGTCCAGAGAAATTCTCTCTCTGTGAGCGGGGTCTGAGCTTGT	584
Db	301	CCAGGGGACGTCCAGAGAAATTCTCTCTCTGTGAGCGGGGTCTGAGCTTGT	360
OY	585	CAAGTTTCATGTGCTGTGAGATGCGTACCTCTCTGAGTCCAGCCAGCCAG	644
Db	361	CAAGTTTCATGTGCTGTGAGATGCGTACCTCTCTGAGTCCAGCCAGCCAG	420
OY	645	TCGCGGCACTGCTCCAGTGCATGTTCCGTACCCCGGCTCAGACCGAGCCAG	704
Db	421	TCGCGGCACTGCTCCAGTGCATGTTCCGTACCCCGGCTCAGACCGAGCCAG	480
OY	705	CCCGCTTCTCCATTCGGAGACCCCTTAAGCCGTAATGAGCCCAACCGGGTATACCAAC	764
Db	481	CCCGCTTCTCCATTCGGAGACCCCTTAAGCCGTAATGAGCCCAACCGGGTATACCAAC	540
OY	765	GAGGCTTCAACAAGCTCTGTGCCAATCTCGCGAGAGATACCCACGGCTCGCTAC	824
Db	541	GAGGCTTCAACAAGCTCTGTGCCAATCTCGCGAGAGATACCCACGGCTCGCTAC	600
OY	825	CCCTTACCCACAGCACTAGGCACTCTCTACCAAGAGCTCCGTTCTACCAAGTTCTCTCC	884
Db	601	CCCTTACCCACAGCACTAGGCACTCTCTACCAAGAGCTCCGTTCTACCAAGTTCTCTCC	660
OY	885	ACCCAGCCGGGGCTGTGTGCTCCGGCAACAACAAGATGTAACGTGTGAACAAGGCCCTTTGG	944
Db	661	ACCCAGCCGGGGCTGTGTGCTCCGGCAACAACAAGATGTAACGTGTGAACAAGGCCCTTTGG	720
OY	945	CTGAATTTTCAACCGGCAACAAACGAGATGATCATCAACAAACAGGAGAGCGCATGTTT	1004
Db	721	CTGAATTTTCAACCGGCAACAAACGAGATGATCATCAACAAACAGGAGAGCGCATGTTT	780
OY	1005	CCTTTTTAAAGTTTTTAACATTTCTGTGCTCGATCCACGGCTCATTTACAATATTTTGTG	1064
Db	781	CCTTTTTAAAGTTTTTAACTTTCTGTGCTCGATCCACGGCTCATTTACAATATTTTGTG	840
OY	1065	GATGTGATTTTTGGCGGATCCCAATCACTGAGAGTTTCAAGAGGCAAAATGGTTCTTTGC	1124
Db	841	GATGTGATTTTTGGCGGATCCCAATCACTGAGAGTTTCAAGAGGCAAAATGGTTCTTTGC	900
OY	1125	GGCAAAAGCGGACCAATGTGAAAGAAATGGGCTCTATATGATCCGGATTCGCCCAAC	1184
Db	901	GGCAAAAGCGGACCAATGTGAAAGAAATGGGCTCTATATGATCCGGATTCGCCCAAC	960
OY	1185	ACTGGGGGCTCACTGTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACGAACAAC	1244
Db	961	ACTGGGGGCTCACTGTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACGAACAAC	1020
OY	1245	AAAGAGCTTCAAAATPACAATGTGGCGAGATGTGTGTTTTCAATCTCTGTGACAAAGTACGAG	1304
Db	1021	AAAGAGCTTCAAAATPACAATGTGGCGAGATGTGTGTTTTCAATCTCTGTGACAAAGTACGAG	1080
OY	1305	CCCGGCTGCAATGTGTGGAATGGAACGAGACCGGACCGGAGGACACTAGCCAGCCCGG	1364
Db	1081	CCCGGCTGCAATGTGTGGAATGGAACGAGACCGGACCGGAGGACACTAGCCAGCCCGG	1144
OY	1365	CGGCTGCAACGTTCACTTTCCCTGAGACTCACTCACTTATGCGCGTCAACGCTTACCAAGAAC	1424
Db	1141	CGGCTGCAACGTTCACTTTCCCTGAGACTCACTCACTTATGCGCGTCAACGCTTACCAAGAAC	1200
OY	1425	ACGGAATTTACACACTGAAAAATAGTACAAACCTTTTGGAAAAAGATTTGGGGATATAT	1484

Db	1201	ACGGATATTACACAACTGAAAATATGATCAACAACCTTTTGCAGAAAGATTTGGGATAT	1260
Qy	1485	TATGACACGATCTTACACCGGCTGTGACATGACCGGCTTGACCCCTCCGCCAACACACTCG	1544
Db	1261	TATGACACGATCTTACACCGGCTGTGACATGACCGGCTTGACCCCTCCGCCAACACACTCG	1320
Qy	1545	CCGGCTCGACGATCGTGCCTCCGGGGCCGCTACCGCATGAGCGGCTCTTTCTTGAGAGAC	1604
Db	1321	CCGGCTCGACGATCGTGCCTCCGGGGCCGCTACCGCATGAGCGGCTCTTTCTTGAGAGAC	1380
Qy	1605	CAGTTCTGTAGCACTACGCCAAGGCCGCTTTCACCCGGGCGCGGCGCGGCGCCCGG	1664
Db	1381	CAGTTCTGTAGCACTACGCCAAGGCCGCTTTCACCCGGGCGCGGCGCGGCGCCCGG	1440
Qy	1665	CCGGGTACCGGACCGGACGCTGCCGACACCAACCGGCTGCTTCCGCGGAGAGAGCCGAG	1724
Db	1441	CCGGGTACCGGACCGGACGCTGCCGACACCAACCGGCTGCTTCCGCGGAGAGAGCCGAG	1500
Qy	1725	GACCCGGGGGCGCCCTCGCCGCAACGCTGTTTGTAGCGCGCGCAACACCGGCTGAGC	1784
Db	1501	GACCCGGGGGCGCCCTCGCCGCAACGCTGTTTGTAGCGCGCGCAACACCGGCTGAGC	1560
Qy	1785	TTTCGGGCTCGGCTATGACACCGGCAACGCTTTCGCGGCGCAACCGGCTGCTC	1844
Db	1561	TTTCGGGCTCGGCTATGACACCGGCAACGCTTTCGCGGCGCAACCGGCTGCTC	1620
Qy	1845	TTTCACCGCGCGCGCGGCTGAAAGCGCTGCCGCTGACAGGCTGACGCTGACCTGCGCGC	1904
Db	1621	TTTCACCGCGCGCGCGGCTGAAAGCGCTGCCGCTGACAGGCTGACGCTGACCTGCGCGC	1680
Qy	1905	CCGCTCGGCTACTACGCGCGGACCGGCTGAGGCGCGCGGCGCGGCGGAGTACGTC	1964
Db	1681	CCGCTCGGCTACTACGCGCGGACCGGCTGAGGCGCGCGGCGCGGCGGAGTACGTC	1740
Qy	1965	GGCACCAGTCCGGCTCGGCTGCTGCCCTGCTGAGCCCAACAGCGCGCGCGC	2017
Db	1741	GGCACCAGTCCGGCTCGGCTGCTGCCCTGCTGAGCCCAACAGCGCGCGCGC	1793
RESULT 7			
BC029289		1793 bp	mRNA linear PRI 31-MAR-2005
LOCUS			
DEFINITION		1793 bp	mRNA linear PRI 31-MAR-2005
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			

Qy	1201	ACGGATATTACACAACTGAAAATATGATCAACAACCTTTTGCAGAAAGATTTGGGATAT	1260
Db	1485	TATGACACGATCTTACACCGGCTGTGACATGACCGGCTTGACCCCTCCGCCAACACACTCG	1544
Qy	1261	TATGACACGATCTTACACCGGCTGTGACATGACCGGCTTGACCCCTCCGCCAACACACTCG	1320
Db	1545	CCGGCTCGACGATCGTGCCTCCGGGGCCGCTACCGCATGAGCGGCTCTTTCTTGAGAGAC	1604
Qy	1321	CCGGCTCGACGATCGTGCCTCCGGGGCCGCTACCGCATGAGCGGCTCTTTCTTGAGAGAC	1380
Db	1605	CAGTTCTGTAGCACTACGCCAAGGCCGCTTTCACCCGGGCGCGGCGCGGCGCCCGG	1664
Qy	1381	CAGTTCTGTAGCACTACGCCAAGGCCGCTTTCACCCGGGCGCGGCGCGGCGCCCGG	1440
Db	1665	CCGGGTACCGGACCGGACGCTGCCGACACCAACCGGCTGCTTCCGCGGAGAGAGCCGAG	1724
Qy	1441	CCGGGTACCGGACCGGACGCTGCCGACACCAACCGGCTGCTTCCGCGGAGAGAGCCGAG	1500
Db	1725	GACCCGGGGGCGCCCTCGCCGCAACGCTGTTTGTAGCGCGCGCAACACCGGCTGAGC	1784
Qy	1501	GACCCGGGGGCGCCCTCGCCGCAACGCTGTTTGTAGCGCGCGCAACACCGGCTGAGC	1560
Db	1785	TTTCGGGCTCGGCTATGACACCGGCAACGCTTTCGCGGCGCAACCGGCTGCTC	1844
Qy	1561	TTTCGGGCTCGGCTATGACACCGGCAACGCTTTCGCGGCGCAACCGGCTGCTC	1620
Db	1845	TTTCACCGCGCGCGCGGCTGAAAGCGCTGCCGCTGACAGGCTGACGCTGACCTGCGCGC	1904
Qy	1621	TTTCACCGCGCGCGCGGCTGAAAGCGCTGCCGCTGACAGGCTGACGCTGACCTGCGCGC	1680
Db	1905	CCGCTCGGCTACTACGCGCGGACCGGCTGAGGCGCGCGGCGCGGCGGAGTACGTC	1964
Qy	1681	CCGCTCGGCTACTACGCGCGGACCGGCTGAGGCGCGCGGCGCGGCGGAGTACGTC	1740
Db	1965	GGCACCAGTCCGGCTCGGCTGCTGCCCTGCTGAGCCCAACAGCGCGCGCGC	2017
Qy	1741	GGCACCAGTCCGGCTCGGCTGCTGCCCTGCTGAGCCCAACAGCGCGCGCGC	1793
FEATURES			
source			
gene			
CDS			
ORIGIN			
Query Match			
Best local similarity			
Matches 1793;			
Conservative			
Mismatches			
Indels			
Gaps			
225			
1			


```

Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatus, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Koyar, C., Kovis, A., Kovis, C., Lara, P.,
Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, K., Liu, Y., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorenshewa, L., Lozada, R., Luk, T., Madu, R.,
Maleshwar, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClalland, H., McPherson, J., Mercado, C., Metzger, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bastaghn, E.,
Noth, A., Nwaokelemeh, O., Obregon, M., Ochi-Okorie, C., Oden, E.,
Okunolu, G., Okunolu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Pirmus, E., Pu, L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savary, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Umami, K., Vargo, C., Verduzco, D., Villasana, D., Vitk, D.,
Voikov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, D.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 101040)
Worley, K. C.
Direct Submission
Submitted (17-JUN-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
-----
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
-----
Project Information
Center project name: FYTU
Center clone name: CH240-161A23
-----
Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 98251 bases at least Q40
Consensus quality: 103096 bases at least Q30
Consensus quality: 106131 bases at least Q20
Estimated insert size: 124062; sum-of-contigs estimation
Estimated insert size: 107267; agarose-fp estimation
Quality coverage: 2x in Q20 bases; agarose-fp estimation
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 4966: contig of 4966 bp in length
* 4967
* 5066: gap of unknown length
* 5067
* 5067: contig of 2515 bp in length
* 5068
* 7681: gap of unknown length
* 7682
* 11922: contig of 4241 bp in length
* 11923
* 12022: gap of unknown length
* 12023
* 16702: contig of 4680 bp in length
* 16703
* 16802: gap of unknown length
* 16803
* 18881: contig of 2079 bp in length
* 18882
* 18981: gap of unknown length
* 22999: contig of 4018 bp in length

```

```

* 23000 23099: gap of unknown length
* 23100 25642: contig of 2543 bp in length
* 25643 25742: gap of unknown length
* 25743 31586: contig of 5844 bp in length
* 31587 31686: gap of unknown length
* 31687 34315: contig of 2629 bp in length
* 34316 34415: gap of unknown length
* 34416 37369: contig of 2954 bp in length
* 37370 37370: gap of unknown length
* 37470 39730: contig of 2261 bp in length
* 39731 39830: gap of unknown length
* 39831 42816: contig of 2986 bp in length
* 42817 42916: gap of unknown length
* 42917 48470: contig of 5554 bp in length
* 48471 48570: gap of unknown length
* 48571 51440: contig of 2870 bp in length
* 51441 51540: gap of unknown length
* 51541 53825: contig of 2285 bp in length
* 53826 53925: gap of unknown length
* 53926 56237: contig of 2312 bp in length
* 56238 56337: gap of unknown length
* 56338 59002: contig of 2665 bp in length
* 59003 59102: gap of unknown length
* 59103 67716: contig of 8614 bp in length
* 67717 67816: gap of unknown length
* 67817 71710: contig of 3894 bp in length
* 71711 71810: gap of unknown length
* 71811 74047: contig of 2237 bp in length
* 74048 74147: gap of unknown length
* 74148 77776: contig of 3629 bp in length
* 77777 77876: gap of unknown length
* 77877 86904: contig of 9028 bp in length
* 86905 87004: gap of unknown length
* 87005 89555: contig of 2551 bp in length
* 89556 89556: gap of unknown length
* 89556 91999: contig of 2344 bp in length
* 92000 92099: gap of unknown length
* 92100 95461: contig of 3362 bp in length
* 95462 95561: gap of unknown length
* 95562 97821: contig of 2260 bp in length
* 97822 97921: gap of unknown length
* 97922 101040: contig of 3119 bp in length.
*
* Location/Qualifiers
* 1..101040
* /organism="Bos taurus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9913"
* /clone="CH240-161A23"
* 4967..5066
* /estimated_length=unknown
* 7582..7681
* /estimated_length=unknown
* 11923..12022
* /estimated_length=unknown
* 16703..16802
* /estimated_length=unknown
* 18882..18981
* /estimated_length=unknown
* 23000..23099
* /estimated_length=unknown
* 25643..25742
* /estimated_length=unknown
* 31587..31686
* /estimated_length=unknown
* 34316..34415
* /estimated_length=unknown
* 37370..37469
* /estimated_length=unknown
* 39731..39830
* /estimated_length=unknown
* 42817..42916
* /estimated_length=unknown
* 48471..48570

```

```
gap /estimated_length=unknown
gap 51441..51540
gap /estimated_length=unknown
gap 53826..53925
gap /estimated_length=unknown
gap 56238..56337
gap /estimated_length=unknown
gap 59003..59102
gap /estimated_length=unknown
gap 67717..67816
gap /estimated_length=unknown
gap 71711..71810
gap /estimated_length=unknown
gap 74048..74147
gap /estimated_length=unknown
gap 77777..77876
gap /estimated_length=unknown
gap 86905..87004
gap /estimated_length=unknown
gap 89556..89655
gap /estimated_length=unknown
gap 92000..92099
gap /estimated_length=unknown
gap 95462..95561
gap /estimated_length=unknown

Query_Match 40.5%; Score 1289; DB 14; Length 101040;
Best Local Similarity 88.0%; Pred. No. 7,68-157;
Matches 1517; Conservative 0; Mismatches 170; Indels 36; Gaps 9;

QY 1490 CACGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTCGCCCAAGCATCGCCGG 1549
DB 68037 CAGGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTCGCCCAAGCATCGCCGG 68096
QY 1550 CTGCGAGATCTGTCGCCGGGGCCCGCTACGCCCATGCGCGGCTCTTCTCGAGAGCAAGTT 1609
DB 68097 CTGCGAGATCTGTCGCCGGGGCCCGCTACGCCCATGCGCGGCTCTTCTCGAGAGCAAGTT 68156
QY 1610 CGTGAAGCACTACGCCAAGGCCCGCTTCCACCGGGCGGGCGCGGGCCCGGGCCGG 1669
DB 68157 CGTGAAGCACTACGCCAAGGCCCGCTTCCACCGGGCGGGCGCGGGCCCGGGCCGG 68216
QY 1670 TACGAGCCGACGCTGCGGCAACCAACGAGGCTGCTGTCGCGGACAGAGCGGAGAGCC 1729
DB 68217 CACGAGCCGACGCTGCGGCAACCAACGAGGCTGCTGTCGCGGACAGAGCGGAGAGCC 68276
QY 1730 GGGCGGCGCTCGCGGCAAGCTGTTTGTGACGCCGGCAACCGGCTGGAATTCCG 1789
DB 68277 GGGCGGCGCTCGCGGCAAGCTGTTTGTGACGCCGGCAACCGGCTGGAATTCCG 68336
QY 1790 GGGCTGGGCTATGAGACGCGGCAAGCTTCCGCGGCAAGCGGCGGCTGCTCTTA 1849
DB 68337 GGGCTGGGCTATGAGACGCGGCAAGCTTCCGCGGCAAGCGGCGGCTGCTCTTA 68396
QY 1850 CGGCGGCGGCGGCTGTAAGCGCTGCGGCTGCGAGGCTGACGCTGACCTGAGCCGCGCT 1909
DB 68397 CGGCGGCGGCGGCTGTAAGCGCTGCGGCTGCGAGGCGGCTGACGCTGAGCCGCGCT 68456
QY 1910 CGGCTACTAGCGGACCCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1969
DB 68457 CGGCTACTAGCGGACCCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 68516
QY 1970 CAGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2029
DB 68517 CAGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 68576
QY 2030 CGGCGGCAATCTCTACTGAGCGAGAGGCGGAGGCGGCTGAGCGGCGGCGGCTGCGCT 2089
DB 68577 CGGCGGCAATCTCTACTGAGCGAGAGGCGGAGGCGGCTGAGCGGCGGCGGCTGCGCT 68636
QY 2090 GCGGCGCGGCGGCGGCGGAGAGCGCAAGCGGCAAGAGACTGTCCGATTCCAGCTGAGTCA 2149
DB 68637 GCGGCGCGGCGGCGGCGGAGAGCGCAAGCGGCAAGAGACTGTCCGATTCCAGCTGAGTCA 68696
```

```
QY 2150 GAGCGCTCTCTGATCAATGCTCATGCACTCCAGCGACTCGGGGATTTACGAGAGGCCAA 2209
DB 68697 GAGCGCTCTCTGATCAATGCTCATGCACTCCAGCGACTCGGGGATTTACGAGAGGCCAA 68756
QY 2210 GCGGAGCGGATCTCGCGGCGGCAACGCGGCTGTCGAGAGTTGTCCTCCGCTCAAGAG 2269
DB 68757 ACGGAGCGGATCTCGCGGCGGCAACGCGGCTGTCGAGAGTTGTCCTCCGCTCAAGAG 68816
QY 2270 CGAGTCTGCGGCGGAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2329
DB 68817 CGAGTCTGCGGCGGAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 68876
QY 2330 TGGCTTTTCTACCGACAGCTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2386
DB 68877 CGGCTTTTCTACCGACAGCTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 68936
QY 2387 ---ACCCAGCGAGCCCTCAGAGCTCTTCCCGAGCTCGGCTCCGACACTCTCTCTT 2443
DB 68937 CCGGCGGCTGCGGCGGCTCTGCGGCGGCTCTGCGGCGGCTCTGCGGCGGCTCTGCGG 68996
QY 2444 GCGCAACCACTCATTTATTTGACCTCTGATGCGGCTGCGAGAGAGAGAGAGAGAGAG 2503
DB 68997 GCGCACTCCCTCATTTATTTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69056
QY 2504 CC-GAGGCTATTTTAA-----CTTTTTCAGAGAGCTCTGCAATT 2548
DB 69057 CCAAGGCTATTTTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 69116
QY 2549 GCTACCGAGCTTCAACTTCTGTAAC-TTTTGTTTTCTTACTTCTTCTTCTG 2607
DB 69117 GCGACCGAGCTTCAACTTCTGTAACCTTTTGGTTTCTTACTTCTTCTG 69176
QY 2608 TGAAGTATCTT-CTTACATTTCCCTTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 2666
DB 69177 TGAAGTATCTTACATTTCCCTTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 69236
QY 2667 TTTCTTATATGAATCTTTCACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2726
DB 69237 TTTCTATATGAATCTTTCACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69296
QY 2727 TTTCCAGCCGCGAAGTCTCGGCTCTCACATTAACATAGATGTTGACTTGAACCTGG 2786
DB 69297 TTTCCAGCCGCGAAGTCTCGGCTCTCACATTAACATAGATGTTGACTTGAACCTGG 69356
QY 2787 ACCCAACGAGCGGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGAT 2846
DB 69357 ACCCAACGAGCGGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATG 69412
QY 2847 ATGTTATATATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2906
DB 69413 ATGTTATATATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69472
QY 2907 GTAAGAGATGCTATTTTATTC---TGTCAAGGCAAAACAGTTCATGCTTAAAC 2963
DB 69473 GTAAGAGATGCTATTTTATTC---TGTCAAGGCAAAACAGTTCATGCTTAAAC 69532
QY 2964 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3023
DB 69533 TTTCT---CATTCATTTCTTCTTCTTCTTCTTCT---CAACTCCCCCGCGGCAAAATGT 69586
QY 3024 TTTTATATTTCTTGTGAGATATATTTTGAAGGCTCTAGAGAGAGAGAGAGAGAGAG 3083
DB 69587 TTTTGTGAGATATTTTGAAGGCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69646
QY 3084 TGAATGAGTTTCCACTTCTCTCAATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 3143
DB 69647 TGAATGAGTTTCTCACTTCTCTCAATCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 69706
QY 3144 CACAAATAGCTTAAGAGAGATCAACCAACCTTTAAGG 3186
DB 69707 CACAAATAGCTTAAGAGAGATCTGATCAACCTTTAAGG 69749
```

RESULT 10
AL845291
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism
Reference
Authors
Title
Journal
Comment

AL845291 209914 bp DNA linear ROD 07-NOV-2002
Mouse DNA sequence from clone RP23-1011 on chromosome 2, complete sequence.
AL845291 GI:23954047
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciuronomachi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 209914)
Blakey, S.
Direct Submission
Submitted (10-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Oct 13, 2002 this sequence version replaced gi:23954047.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (1.e., phred quality >= 30) or an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-1011 is from the RP23-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see <http://www.choi.org/bacpac/home.htm>
VECTOR: pBACe3.6.

FEATURES
source
1. 209914
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-1011"
/clone_lib="RP23-23"

ORIGIN
Query Match 36.6%; Score 1164.6; DB 9; Length 209914;
Best Local Similarity 85.6%; Pred. No. 7.1e-141;
Matches 1994; Conservative 0; Mismatches 179; Indels 72; Gaps 15;

1490 CACGATCTACACCGGTGTGATGATGACCGCTGACCCCTGCGCAAGACGATCGCGCG 1549
186044 CAGGATCTACACGCGGTGTGATGATGACCGCTGACCCCTGCGCAAGACGATCTCGCG 186103
1550 CTCGAGATCTGCG 1609
186104 CTCGAGATCTGCG 186163
1610 CGTGAGCACTACGCGCAAGCGCGCTTCACCCGCGCGCGCGCGCGCGCGCGCGCGCG 1669

186164 CGTGAGCACTACGCGCAAGCGCGCTTCACCCGCGCGCGCGCGCGCGCGCGCGCGCG 186223
1670 TACGAGACCGGACGCTGTGCGGACACCAACGAGGCTGTCTCCCGGAGAGCGGAGACCC 1729
186224 CACGAGACCGGACGCTGTGCGGACACCAACGAGGCTGTCTCCCGGAGAGCGGAGACCC 186283
1730 GGGCG 1789
186284 GGGCG 186343
1790 GGGCTGCGCGCTTATGACACGCGGACGACTTTCGCGGCAACGCGGCGACGCTGTCTTGA 1849
186344 GGGCTGCGCGCTTATGACACGCGGACGACTTTCGCGGCAACGCGGCGACGCTGTCTTGA 186403
1850 GCG 1909
186404 TCG 186463
1910 CGGCTACTACGCGCGGACCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1969
186464 CGGCTACTACGCGGACCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186523
1970 CAAGTCGCGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029
186524 CAAGTCGCGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186583
2030 CGGCGCGCAATCCCTACTGCGGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 2089
186584 CGGCGCGCAATCCCTACTGCGGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 186643
2090 GCG 2149
186644 GCG 186700
2150 GAGCGCGCTCTGATCAAGTCATGCACTCCAGCGACTCGGCGGATTTTCAAGAGCGCAA 2209
186701 GAGCGCGCTCTGATCAAGTCATGCACTCCAGCGACTCGGCGGATTTTCAAGAGCGCAA 186760
2210 GCGGAGCGGATCTGCG 2269
186761 GCGGAGCGGATCTGCG 186820
2270 CGAGGCTCTGCG 2329
186821 CGAGGCTCTGCG 186880
2330 TGGCTTCTACTGCGGACGCTTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2389
186881 TGGCTTCTACTGCGGACGCTTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186940
2390 CCGAGCGAGCG 2449
186941 CCGAGCGAGCG 186988
2450 CCACTCATTTTATTTGACCGCTCGATGCGCGCTCTGCAAGCAATAGTCAGGTCGCGAGC 2509
186988 TCGGCTC-TTTCACCTTGAACCTCGATGACGCTGCTGCG-GGGATTAAGTCAAGTCTCTACT 187046
2510 GTGATTTTAACTTTT-----TGCACGACGCTCTTGCAC 2545
187047 ATGATTTTAACTTTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 187106
2546 TTACTCATCGGACCTTGAACCTTTGCTGCTTAAACCTTTTGTGTTTCTTACTTCTTCTTC 2605
187107 TTACTCATCGGACCTTGAACCTTTGCTGCTTAAACCTTTGTTGTTTCTTCAACTT---TTCGTC 187162
2606 TGTGAGTTATCTCTTAC-----AATCCCGCTCCCGCTGCTTCTTTCTTTCTTTCTTT 2659
187163 TGTGAGTTATGATCTCTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 187222
2660 CTCTCTCTTCTTGTATGAATCTTTTCACTTTTGAAGACCTGCGGCACTCTGTGAGGCA 2719
187223 CTCTCTTCTCTTGAATGAATCTTTTCACTTTTGAAGACCTGCGGCACTCTGTGAGGCA 187282

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 176172: contig of 176172 bp in length.

FEATURES

source

1. 176172
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-324F12"
3093..5080
/note="wgs end extension
clone end:Sp6"
5131..6045
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ240781"
124466..126024
/note="wgs contig"
127084..128675
/note="wgs_contig"
complement(173779..174717)
/note="clone_boundary
clone_end:T7
site:
end_sequence:BZ240185"

ORIGIN

Query Match 36.3%; Score 1157.8; DB 14; Length 176172;
Best Local Similarity 84.9%; Pred. No. 5,5e-140;
Matches 1471; Conservative 0; Mismatches 197; Indels 65; Gaps 13;

QY 1490 CACGATCTTACACCGGCTGTGACATGAGACCGCTGACCCCTCGCCCAAGACTCCGCCG 1549
Db 88390 CAGGATCTTACACGCGGCTGTGACATGAGACCGCTTGAACCCGCTGCCCAAGACTCTCCG 88331
QY 1550 CTGCGAGATCTGACCGCGGCGCGCTTACGCGATGCGCGGCTCTTCTCTGACAGACCAT 1609
Db 88330 CTCTAGATCTGCGCGCGCGCTCTGCTACCGCATGCGCGGCTCTTCTCTGACAGATCAAT 88271
QY 1610 CGTGAACAATAACGCCAAGCCCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 1669
Db 88270 CGTGAACAATAACGCCAAGCCCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 88211
QY 1670 TACGACCGCGACCGCTGCGCAACCAACGAGGCTGCTGTGCGCGCGCAACCGGCTGGA 1729
Db 88210 CACGACCGCGACCGCTGCGCAACCAACGAGGCTTACTGTCCCGCAACGAGCGGAGAGACC 88151
QY 1730 GGGCGCGCGCTGCGCGCAACCGCTGTTGTGAAGCGCGCGCAACCGGCTGGA 1789
Db 88150 GGGCGCGCGCTGCGCGCAACCGCTGTTGTGAAGCGCGCGCAACCGGCTGGA 88091
QY 1790 GGCCTCGGCTATGACACCGGCAACCGACTTTCGCGGCAACCGGCGCAACCGCTCTCT 1849
Db 88090 GGCCTCGGCTATGACACCGGCAACCGACTTTCGCGGCAACCGGCGCAACCGCTCTCT 88031
QY 1850 CGCGCGCGCGCGCGCTGGAAGCGCTGCGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1909
Db 88030 CGCGCGCGCGCGCGCTGGAAGCGCTGCGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 87971
QY 1910 CGGCTACTAGCGCGACCGGCTCGGCGCTGAGGCGCGCGCGAGTCCCGCGAGTACTGAGG 1969
Db 87970 CGGCTACTAGCGCGACCGGCTCGGCGCTGAGGCGCGCGCGAGTCCCGCGAGTACTGAG 87911
QY 1970 CAAGTGGGCTCGGTGCTGCGCTGCTGCGCGCAACGCGCGCGCGCGCGCGCGATGAG 2029
|||||

Db 87910 CAAGTGGGCTCGGTGCTGCGCTGCTGCGCGCAACGCGCTGCGCGCGCGCGCGATGAG 87851
QY 2030 CGGCGCGCAATCCCTTACCTGCGCGGAGCGCGAGGCGCTGCGCGCGAGGCTCGCGCGCT 2089
Db 87850 GGGCGCGCAACCTTACTTAAAGCGGAGGCGGAGGCGCTGCGCGCGAGGCTCTCGCGCT 87791
QY 2090 GCCCGCGCGCGCGCGGAGGAGCGCAAGGCGCAAGGCTGCGCGATTCAGCTGATGCA 2149
Db 87790 GCGCGC---CGCGCGAGGAGCGCGCAAGGCGCAAGGCTTCCGACTCGAGCTGATGCA 87734
QY 2150 GAGCGCTCTCGATCAATGATTCATCGACTCGAGGCACTCGGAGATTTTACAGAGCGCAA 2209
Db 87733 GAGCGCTCTCGATCAATGATTCATCGACTCGAGGCACTCGGAGATTTTACAGAGCGCAA 87674
QY 2210 GCGAGGCGGATCTTCGCGCGCGGCAAGCGCGGCTCGAGAGTTCGCGCGCGCTCAAGAG 2269
Db 87673 GCGAGGCGGATCTTCGCGCGCGGCAAGCGCGGCTCGAGAGTTCGCGCGCGCTCAAGAG 87614
QY 2270 CGAGGTCTGCGCGCGGAGCTGCGAGAGAACTGCGCGCAAGGACATTAAGCGCTACTA 2329
Db 87613 CGAGGTCTGCGCGCGGAGCTGCGAGAGAACTGCGCGCAAGGACATTAAGCGCTACTA 87554
QY 2330 TGGCTTACTCGGACAGCTAAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGAGC 2389
Db 87553 TGGCTTACTCGGACAGCTAAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGAGC 87504
QY 2390 CCCAGCGAGCGCGCTCAAGGCTTCCCGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCG 2449
Db 87503 -----GCGCGTACGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 87454
QY 2450 CCAGCTATTTATTTATTTGACCGCTCGATGCGCGCTGCGAGCGAATAGTCAAGTCTCGAGC 2509
Db 87453 TCGCTC--TTTACCTTGAAGCTCGATGACCGCTGCG--GGGATTAAGTCAAGTCTCCAC 87396
QY 2510 GTGATTTTAACTTTT-----TGACAGAGCTCTGCAATTAGCTCA 2553
Db 87395 ATGATTTTAAACTTTT-----TGCTTCTTCTGCGAGCGCACTCTGCGAGTTAGGCA 87336
QY 2554 CGGACTTTAACTTTGCGCTGTAACCTTTGGTTTCTTACTCTCTCTCTGCGAGT 2613
Db 87335 CGGACTTTAACTTTGCGCTGTAACCTTTGGTTTCTTACTCTCTCTCTGCGAGT 87280
QY 2614 TATCTCTTAC-----AATTCCTTCCCGCGCGCTCTTCTTCTTACTCTCTCTCT 2667
Db 87279 TATGATCTCTCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 87220
QY 2668 TCTTGTATGAATCTTTTCACTTTAGAGACCTGCGAGTCTGCTGAGGAGCGAGT 2727
Db 87219 TCTTGTATGAATCTTTTCACTTTAGAGACCTGCGAGTCTGCTGAGGAGCGAGT 87160
QY 2728 TCGGACCGCGCAAGCTTCGCGCTTCAATTAACATAGAGTGTGAATCTTGAACCTGGA 2787
Db 87159 TCGGACCGCGCAAGCTTCGCGCTTCAATTAACATAGAGTGTGAATCTTGAACCTGGA 87100
QY 2788 CCCAGCGAGCGCTCTTCTTATCCCGAGTGTAGTGAATGATGATGATGATGATGATGATGAT 2847
Db 87099 CCCAGCGAGCGCTCTTCTTATCCCGAGTGTAGTGAATGATGATGATGATGATGATGAT 87040
QY 2848 TGTTAATTAATTTAGTGAACAAAGCTGTGAATGATGATGATGATGATGATGATGATGAT 2907
Db 87039 TGTTAATTAATTTAGTGAACAAAGCTGTGAATGATGATGATGATGATGATGATGATGAT 86980
QY 2908 TAAAGATGCGTATTTTATTTTTC---TCGTCAAGGCGCAAAACAGTCAAGCTTAACCT 2964
Db 86979 TAAAGATGCGTATTTTATTTTTC---TCGTCAAGGCGCAAAACAGTCAAGCTTAACCT 86920
QY 2965 TTTTCT 3020
Db 86919 TTTATACCT 86860
QY 3021 CTCTTTTAAATTTCTTGTGAGATA-----ATATTTAAGAGGCTTGAAGACATGAAT 3074
Db 86859 CCGCTTTGTTTCTTGTGAGATTTTAAAGATTTTAAAGAGGCTTGAAGACATGAAT 86800

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
OY	3075	ACTGAGATGATATGGTGTCCCATCTTCCCTCCCAATCCGTTGACGAATATATTTCTATGT	AC165777/c				
DB	86799	AAACCTTATAGTGTGGCTTCTTCACCTTCCAGTGTGTTGACATGATACCACTATGT					
OY	3135	GCCCTAATGCACCA-AATAGCTAAGAGAAATCCACCAACACCTTTAAAG	AC165777				
DB	86739	GCCCCAGTGCACACAGCATCTAGTAAGAGAAATCCACCTGAAACATCTATAAAG					
RESULT 12							
AC165777		101040 bp	DNA	linear	HTG 17-JUN-2005		
Bos taurus clone CH240-161A23, *** SEQUENCING IN PROGRESS ***							
unordered pieces.							
AC165777							
AC165777.1 GI:70912797							
HTG, HTGS, PHASE1.							
Bos taurus (cow)							
Bos taurus							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;							
Bovidae; Bovinae; Bos.							
1 (bases 1 to 101040)							
Munry, D., Adams, C., Agpai II, O., Allen, C., Albrooks, S., Archer, P.,							
Atencio, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,							
Bereducio, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,							
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,							
Chacko, J., Chahour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,							
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,							
Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De							
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,							
Dzudin, J., McCanley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,							
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,							
Gaskin, C., Gao, Q., Fu, E., Garcia, A., Garcia, R., Garner, I.,							
Gonzalez-Garay, M., Guevara, K., Holder, M., Haaland, W., Heberlein, K.,							
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Hawlak, P.,							
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,							
Hines, S., Hitchens, M., Hodgson, A., Hughes, M., Holline, B.,							
Howell, L. T., Huylk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,							
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatus, K., Kelly, S.,							
Keys, T., Khan, Z., King, L., Koyar, C., Kowis, A., Kowis, C., Lara, F.,							
Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,							
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Llyanage, D.,							
London, P., Lopez, J., Lorenuhewa, L., Losado, R., Luk, T., Madu, R.,							
Maneshahri, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,							
McCluskey, H., McPherson, J., Mercedao, C., Metzker, M.,							
Miloslavjevic, A., Minja, E., Morgan, M., Morris, S., Mundasa, M.,							
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Easaugh, E.,							
Noth, A., Nwaelelemeh, O., Obregon, M., Och-Olorie, C., Odeh, E.,							
Okwomun, G., Okwomun, K., Parker, D., Pasternak, S., Patel, B.,							
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, V., Pham, T.,							
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quintoz, J.,							
Rabeta, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,							
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sama, M., Sanders, W.,							
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,							
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,							
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,							
Trejos, Z., Uemami, K., Varago, C., Verdusco, D., Villasana, D., Virk, D.,							
Velkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, D.,							
Welch, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,							
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,			</				

COMMENT

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu

----- Project Information
Center project name: FJUJ
Center Clone name: CH240-161A23

----- Summary Statistics
Sequencing vector: Plasmid,
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 98251 bases at least Q40
Consensus quality: 103096 bases at least Q30
Consensus quality: 106131 bases at least Q20
Estimated insert size: 124062; sum-of-contigs estimation
Estimated insert size: 107277; agarose-fp estimation
Quality coverage: 2x in Q20 bases; agarose-fp estimation
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_date.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4966: contig of 4966 bp in length
4967 5066: gap of unknown length
5067 7581: contig of 2515 bp in length
7582 7681: gap of unknown length
7682 11922: contig of 4241 bp in length
11923 12023: gap of unknown length
12023 16702: contig of 4680 bp in length
16703 16802: gap of unknown length
16803 18881: gap of 2079 bp in length
18882 18981: gap of unknown length
18982 22999: contig of 4018 bp in length
23000 23099: gap of unknown length
23100 25642: contig of 2543 bp in length
25643 25742: gap of unknown length
25743 31586: contig of 5844 bp in length
31587 31686: gap of unknown length
31687 34315: contig of 2629 bp in length
34316 34415: gap of unknown length
34416 37369: contig of 2934 bp in length
37370 37469: gap of unknown length
37470 39730: contig of 2261 bp in length
39731 39830: gap of unknown length
39831 42816: contig of 2986 bp in length
42817 42916: gap of unknown length
42917 48470: contig of 5554 bp in length
48471 48570: gap of unknown length
48571 51440: contig of 2870 bp in length
51441 51540: gap of unknown length
51541 53825: contig of 2285 bp in length
53826 53925: gap of unknown length
53926 56237: contig of 2312 bp in length
56238 56337: gap of unknown length
56338 59002: contig of 2665 bp in length
59003 59102: gap of unknown length
59103 67716: contig of 8614 bp in length
67717 67816: gap of unknown length
67817 71710: contig of 3894 bp in length
71711 71810: gap of unknown length
71811 74047: contig of 2237 bp in length
74048 74147: gap of unknown length
74148 77776: contig of 3629 bp in length
77777 77876: gap of unknown length
77877 86904: contig of 9028 bp in length

```

```
* 86905 87004: gap of unknown length
* 87005 89555: contig of 2551 bp in length
* 89556 89555: gap of unknown length
* 89556 91959: contig of 2344 bp in length
* 92000 92099: gap of unknown length
* 92100 95461: contig of 3362 bp in length
* 95462 95561: gap of unknown length
* 95562 97821: contig of 2260 bp in length
* 97822 97921: gap of unknown length
* 97922 101040: contig of 3119 bp in length.
Location/Qualifiers
1..101040
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/cldone="CH240-161A23"
4967..5066
/estimated_length=unknown
7582..7681
/estimated_length=unknown
11923..12022
/estimated_length=unknown
16703..16802
/estimated_length=unknown
18882..18981
/estimated_length=unknown
23000..23099
/estimated_length=unknown
25643..25742
/estimated_length=unknown
31587..31686
/estimated_length=unknown
34316..34415
/estimated_length=unknown
37370..37469
/estimated_length=unknown
39731..39830
/estimated_length=unknown
42817..42916
/estimated_length=unknown
48471..48570
/estimated_length=unknown
51441..51540
/estimated_length=unknown
53826..53925
/estimated_length=unknown
56238..56337
/estimated_length=unknown
59003..59102
/estimated_length=unknown
67717..67816
/estimated_length=unknown
71711..71810
/estimated_length=unknown
74048..74147
/estimated_length=unknown
77777..77876
/estimated_length=unknown
86905..87004
/estimated_length=unknown
89556..89655
/estimated_length=unknown
92000..92099
/estimated_length=unknown
95462..95561
/estimated_length=unknown
gap
/estimated_length=unknown
Query Match 30.5%; Score 971; DB 14; Length 101040;
Best Local Similarity 98.5%; Pred. No. 6,7e-116;
Matches 980; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
Dh 65426 CAGGTGATTAATCTAATTAATGTCATCTAATTAATTAATTAATGTCAGAGTTAAACCAATGG 65367
Qy 61 CAGAGCCGTTTCAATGCGGTGCAACAGACCAAGATCAAAAGAGCCTTTCTGATGTC 120
Dh 65366 CAGAGCGTTTCAATGCGGTGCAACAGACCAAGATCAAAAGAGCCTTTCTGATGTC 65307
Qy 121 TGCATATGTCATATTGGCCATCTCTTCTCCAGGAGAAAAAGTAATCAAACTTT 180
Dh 65306 TGCATATGTCATATTGGCCATCTCTTCTCCAGGAGAAAAAGTAATCAAACTTT 65247
Qy 181 TGAGAACATTTTGTGTTGAAGTCTTTCTGTCTAGTGAAGGGGCTTGAGATTTCTAG 240
Dh 65246 TGAGAACATTTTGTGTTGAAGTCTTTCTGTCTAGTGAAGGGGCTTGAGATTTCTAG 65187
Qy 241 TTTATGATTAATAGGACTTTAAAAACAGGAGACAGAGGCGAGTGTTCAGTTCTAGAG 300
Dh 65186 TTTATGATTAATAGGACTTTAAAAACAGGAGACAGAGGCGAGTGTTCAGTTCTAGAG 65127
Qy 301 CTATGACGCTGAGACACTGCTTTCTTCTATCATGCTCTCCAGAAATTTCTCAATG 360
Dh 65126 CTATGACGCTGAGACACTGCTTTCTTCTATCATGCTCTCCAGAAATTTCTCAATG 65067
Qy 361 TGAGCAGAGCTACCACTTCAAGCGGATCCAGCTTGTTCGACGATCATCCATT 420
Dh 65066 TGAGCAGAGCTACCACTTCAAGCGGATCCAGCTTGTTCGACGATCATCCATT 65007
Qy 421 TCTGACACCTGACCAACTGAGAGAGTCACTTTGAAAAAATTAACAGGAGGAGTGA 480
Dh 65006 TCTGACACCTGACCAACTGAGAGAGTCACTTTGAAAAAATTAACAGGAGGAGTGA 64947
Qy 481 CGAATCAGTCAATACAGACAAATTTCTGACTCAAGAGACTCACAGGAGGAGTCCAGA 540
Dh 64946 CGAATCAGTCAATACAGACAAATTTCTGACTCAAGAGACTCACAGGAGGAGTCCAGA 64887
Qy 541 GAAATTAATCTCTCTCTGTTGAGCGGGTCTCTGAGCTTCTGACAGTTTGATGGCT 600
Dh 64886 GAAATTAATCTCTCTCTGTTGAGCGGGTCTCTGAGCTTCTGACAGTTTGATGGCT 64827
Qy 601 CTGCTGAGATCGGTACTCTCTCTCAATCGAGCCAGCCAGTCTGGGGCACTGCTC 660
Dh 64826 CTGCTGAGATCGGTACTCTCTCTCAATCGAGCCAGCCAGTCTGGGGCACTGCTC 64767
Qy 661 CCAATGCCATGTTCCCGTACCCGCGCAGACCGAGCCAGCCCGCTTCTCAATCG 720
Dh 64766 CCAATGCCATGTTCCCGTACCCGCGCAGACCGAGCCAGCCCGCTTCTCAATCG 64707
Qy 721 GCAGCCCTAGCCGCTCATGAGCCCAACCCGGTCAATCAACAGGAGCTTCAACAGCC 780
Dh 64706 GCAGCCCTAGCCGCTCATGAGCCCAACCCGGTCAATCAACAGGAGCTTCAACAGCC 64647
Qy 781 TCCGTGCAACTCTGCGCGAGGAGTACCCACAGCGCGGCTACCCCTACCAAGAGT 840
Dh 64646 TCCGTGCAACTCTGCGCGAGGAGTACCCACAGCGCGGCTACCCCTACCAAGAGT 64587
Qy 841 ACGGCACCTCTACCAAGAGACTCCGTTCAACAGTTCTCTCCACCCAGCCGCGGCTG 900
Dh 64586 ACGGCACCTCTACCAAGAGAGCCCGTTCAACAGTTCTCTCCACCAAGCCGCGGCTG 64527
Qy 901 TGCCCGCAAGCAACAGGTGACTGTCACACAGCCCTTTGGCTGAATTTCAACGGC 960
Dh 64526 TGCCCGCAAGCAACAGGTGACTGTCACACAGCCCTTTGGCTGAATTTCAACGGC 64467
Qy 961 ACCAAAGGAGATGATCATCAACCAAGGAGG 995
Dh 64466 ACCAAAGGAGATGATCATCAACCAAGGAGG 64432
```

RESULT 13
AF287006 2218 bp mRNA linear VRT 10-JAN-2001
LOCUS AF287006
DEFINITION Danio rerio T-box brain 1 mRNA, partial cds.
ACCESSION AF287006
VERSION AF287006.1 GI:12060813

KEYWORDS

Danio rerio (zebrafish)

SOURCE

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 2218)

Mione, M., Shanmugalingam, S., Kimmel, D. and Griffin, K.

Overlapping expression of zebrafish T-brain-1 and eomesodermin

during forebrain development

Mech. Dev. 100 (1), 93-97 (2001)

11118891

2 (bases 1 to 2218)

Mione, M. and Griffin, K.

Direct Submission

Submitted (13-JUL-2000) Anatomy, University College London, Gower

Street, London WC1E6BT, UK

Location/Qualifiers

1..2218

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

<1..1314

/note="tbr"

/codon_start=1

/product="T-box brain 1"

/protein_id="AAG48249.1"

/db_xref="GI:12060814"

/translation="SGLDPTAHYNTFVDVILADPNHMRFGGKAVPCGKADTNVGNR

VTHMDEPNTGAMHMRQELSPKRLTKTNKGNNTNGMNVQSLHKYOPRLAVQNV

EGTDETSQGRVQTEFPETQFIATVATONTDOLKIDHPFAKGFNDYTYTG

CDIDRLTPSPGDSPSQIMPSRVMATGSLFDQFMSYAKSRFHFGVGAAGTQSV

PLSNLSLPQTEETVAPORFVTPANNRLDFAASADAADAFANATLTLSTA

AAGVALPLPGAGCSNRPLGYGDEPGVGRTPQYCKSSSVLSCPSMSVGSRTTAS

SVIVPELEGGDPIAPRSPRPLGAGDESKPDLSBSWIEPSSIKLSDSDSIFQAK

KRRISPAIVSETSSPLKSBMLTPRECKNCSKDIQYKTFPHS"

FEATURES

source

1..2218

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

<1..1314

/note="tbr"

/codon_start=1

/product="T-box brain 1"

/protein_id="AAG48249.1"

/db_xref="GI:12060814"

/translation="SGLDPTAHYNTFVDVILADPNHMRFGGKAVPCGKADTNVGNR

VTHMDEPNTGAMHMRQELSPKRLTKTNKGNNTNGMNVQSLHKYOPRLAVQNV

EGTDETSQGRVQTEFPETQFIATVATONTDOLKIDHPFAKGFNDYTYTG

CDIDRLTPSPGDSPSQIMPSRVMATGSLFDQFMSYAKSRFHFGVGAAGTQSV

PLSNLSLPQTEETVAPORFVTPANNRLDFAASADAADAFANATLTLSTA

AAGVALPLPGAGCSNRPLGYGDEPGVGRTPQYCKSSSVLSCPSMSVGSRTTAS

SVIVPELEGGDPIAPRSPRPLGAGDESKPDLSBSWIEPSSIKLSDSDSIFQAK

KRRISPAIVSETSSPLKSBMLTPRECKNCSKDIQYKTFPHS"

ORIGIN

Query Match

Best Local Similarity 21.8%; Score 694.8; DB 5; Length 2218;

Matches 971; Conservative 0; Mismatches 337; Indels 24; Gaps 5;

1026 TCTGCTCTGCATCCACGGCTCATTAATATTTTGTGATGATTTTGGCGATCC 1085

1 TCTGCTCTGCATCCACGGCTCATTAATATTTTGTGATGATTTTGGCGATCC 60

1086 AATCATCTGAGATTTCAAGAGGCAATGGTCTCTTGGGGCAAGCGACACCATG 1145

61 AATCATCTGAGATTTCAAGAGGCAATGGTCTCTTGGGGCAAGCGACACCATG 120

1146 CAAGAATCTCTTATATGATCCGATTTCCCAACACTGGGCTCATGATGCG 1205

121 ACGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 180

1206 CAAGAATCTCTTATATGATCCGATTTCCCAACACTGGGCTCATGATGCG 1265

181 CAAGAATCTCTTATATGATCCGATTTCCCAACACTGGGCTCATGATGCG 240

1266 GGGCAATGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATG 1325

241 GGAAGATGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATG 300

1326 GTGACAGGAGCGGACGAGACACTAGCCGCGCGCTGACGATTCCTTC 1385

301 GTGACAGGAGCGGACGAGACACTAGCCGCGCGCTGACGATTCCTTC 360

1386 CTTGAGACTGATTCATGCTGCTACCGCTTACGAGACGAGATTTTACAA 1445

361 CCGAGAGCGAGTTTCTTCTGCTACCGCTTACGAGACGAGATTTTACAA 420

1446 ATGATATACAACTTTTGGAAAAGATTTTGGGATTAATATGACACGATC 1505

421 ATGACACCAACCACTTTTGGAAAAGATTTTGGGATTAATATGACACGATC 480

QY 1506 TGTGACATGACCGGCTGACCCCTGCGCCCAACGATCGCGGCTGACGATCGTCC 1565

DB 481 TGTGACATGACCGGCTGACCCCTGCGCCCAACGATCGCGGCTGACGATCGTCC 540

QY 1566 GGGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGAC 1625

DB 541 AGCGACAGATGATGACGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 1626 AAGGCGGCTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1685

DB 601 AAGTCCGCTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 654

QY 1686 CCGACACCAACGCGGCTGCTGCGCGGAGAGGCGGAGGCGGCGGCGGCGGCGG 1745

DB 655 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 714

QY 1746 CAAGCTGTTTGTGACCGCGGCAACGCGGCTGACCTTGGGCGCTGCGCTT 1801

DB 715 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774

QY 1802 --TGACAGGCGGACGATTTGCGGCGGCAACGCGGCGGCGGCGGCGGCGG 1859

DB 775 GCTGCTGCGGCTGATGATTTGCGGCGGCAACGCGGCGGCGGCGGCGGCGG 834

QY 1860 GCGGTGAAGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1919

DB 835 GAGATTAAGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 894

QY 1920 GCCGACCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1979

DB 895 GCGGAGCGG-----GGGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 948

QY 1980 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039

DB 949 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008

QY 2040 CCTTACTGCGGCGGAGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 2099

DB 1009 CTGCTGCTGCGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065

QY 2100 GCGGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2159

DB 1066 GGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1125

QY 2160 TCGATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2219

DB 1126 TCAATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1185

QY 2220 ATCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2279

DB 1186 ATTTGCGGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245

QY 2280 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2339

DB 1246 ACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1302

QY 2340 TCGCAAGCTAG 2351

DB 1303 CCCCAAGTTAG 1314

RESULT 14

AR655897/c AR655897 514 bp DNA linear PAT 13-JUN-2005

LOCUS Sequence 37 from patent US 6893820.

ACCESSION AR655897

VERSION AR655897.1 GI:67588385

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 514)

AUTHORS Plass, C.
TITLE Detection of methylated CpG rich sequences diagnostic for malignant cells
JOURNAL Patent: US 6893820-A 37 17-MAY-2005;
The Ohio State University Research Foundation; Columbus, OH
FEATURES
source
1..514
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 16.1%; Score 514; DB 6; Length 514;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GCTGTGATGATGACCGCTGACCCCTGCGCCCAAGCATCGCGCGCTCGAGATCGTGC 1563
Db 514 GCTGTGATGATGACCGCTGACCCCTGCGCCCAAGCATCGCGCGCTCGAGATCGTGC 455
QY 1564 CCGGGGCGCGCTGACCGCATGGCGGCTCTTTCTGCAAGACCATGTTGGTAGCACTAG 1623
Db 454 CCGGGGCGCGCTGACCGCATGGCGGCTCTTTCTGCAAGACCATGTTGGTAGCACTAG 395
QY 1624 CCAAGGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683
Db 394 CCAAGGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
QY 1684 TGGCGCACACCAACCGGCTGCTGTGCGCGAGACGCGAGAGACCGCGCGCGCGCTGCG 1743
Db 334 TGGCGCACACCAACCGGCTGCTGTGCGCGAGACGCGAGAGACCGCGCGCGCGCTGCG 275
QY 1744 CGCAACGCGTGGTTGTAGCGCGCGCAACACCGGCTGGAATTGGCGCGCTCGGCTATG 1803
Db 274 CGCAACGCGTGGTTGTAGCGCGCGCAACACCGGCTGGAATTGGCGCGCTCGGCTATG 215
QY 1804 ACACGCGCACGACCTTTCGCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1863
Db 214 ACACGCGCACGACCTTTCGCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 155
QY 1864 TGAAGCGCTGCGCTGCGAGAGCTGAGAGCTGACCTGCGCGCGCGCGCGCGCGCGCG 1923
Db 154 TGAAGCGCTGCGCTGCGAGAGCTGAGAGCTGACCTGCGCGCGCGCGCGCGCGCGCG 95
QY 1924 ACCCGTCGGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983
Db 94 ACCCGTCGGGCTGCGGCG 35
QY 1984 TGCTGCTGCTGCGCGCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2017
Db 34 TGCTGCTGCTGCGCGCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1

RESULT 15
HSA341671 793 bp DNA linear PRI 18-JUL-2002
LOCUS HSA341671
DEFINITION Homo sapiens genomic sequence surrounding Not1 site, clone
N16-LP8R.
VERSION AJ341671
ACCESSION AJ341671.1 GI:15886089
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Kuzenko, A.S., Gzazacullin, R.Z., Al-Amin, A.N., Wang, F., Kashba, S.M.,
Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,
Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I.,
Kiselev, L.L., Wasserman, W., Wahlestedt, C., and Zadorovskiy, E.R.
TITLE Not1 flanking sequences: a tool for gene discovery and verification
of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)

PUBMED 12136098
REFERENCE 2 (bases 1 to 793)
AUTHORS Zadorovskiy, E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source
1..793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="N16-LP8R"

ORIGIN

Query Match 15.4%; Score 491.2; DB 8; Length 793;
Best Local Similarity 87.0%; Pred. No. 8.2e-54;
Matches 574; Conservative 0; Mismatches 80; Indels 6; Gaps 3;

QY 2010 GCGGCGCGCGCGCATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2069
Db 1 GCGGCGCGCGCGCATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 2070 GCCGCGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2129
Db 61 GCCGCGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 2130 TCCGATTCCAGCTGATGAGAGCGCCCTCTCGATCAAGTCCATCCACTCCAGACTCG 2189
Db 121 TCCGATTCCAGCTGATGAGAGCGCCCTCTCGATCAAGTCCATCCACTCCAGACTCG 180
QY 2190 GGAATTACAGAGAGGCGCAAGCGGAGCGGATCTGCGCGCGCGCGCGCGCGCGCGCGCG 2249
Db 181 GGAATTACAGAGAGGCGCAAGCGGAGCGGATCTGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 2250 AGTTGCTCCCGCTCAAGAGCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2309
Db 241 AGTTGCTCCCGCTCAAGAGCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 2310 AAGGACATTAGCGGCTACTATGCTTACTGCGACAGCTAGGCGCGCGCGCGCGCGCGCG 2369
Db 301 AAGGACATTAGCGGCTACTATGCTTACTGCGACAGCTAGGCGCGCGCGCGCGCGCGCG 360
QY 2370 GCCCG 2429
Db 361 GCCCG 420
QY 2430 CCAACTCTCTCTTGGCGACCACTCATTTTATTTGAACCTGTGAGGCGGCTTGCAGCGA 2489
Db 421 CCAACTCTCTCTTGGCGACCACTCATTTTATTTGAACCTGTGAGGCGGCTTGCAGCGA 480
QY 2490 ATTAGTGCAGGTCTCCGAG-CGTGATTTTAACTTTTTCAGACAGAGCTGCGCAATTA 2548
Db 481 ATTAGTGCAGGTCTCCGAGCGGTGATATTAACCTTTTTCGAGAGGTCTTGC-ATAA 539
QY 2549 GCTCAGCGACTTCACTTGTGTGTAACCTTTTGTGTTTCTTAATTAATCTTCTTGTGT 2608
Db 540 GGTACAGGCGCTTAATTTTTCGGAACATTTGGGTTCTTC---TTACCTCTTCTT 595
QY 2609 GAGATTATCTCTCAATTCATTCCTCCCTGCTCTTCTTCACTCACTCACTTCTCTT 2668
Db 596 GAGAGAAATATCTCCAAACCGCGCGCGGTTTATTTATTCCTCTCTTCTTCTTCTT 655

Search completed: December 20, 2005, 12:35:54
Job time: 15693 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 03:16:07 ; Search time 1805 Seconds
(without alignments)
11763.829 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....caccacaacaccttaag 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq.21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	3186	100.0	3186	12 ADL97817	ADL97817 Human TBR
2	1793	56.3	1793	11 AEA79130	Aea79130 Human apo
3	1695.4	53.2	6878	5 ABA14582	Abal4582 Human ner
4	864	27.1	864	12 ACH87502	Ach87502 Human gen
5	604	19.0	1026	6 ABO22410	Abq22410 Oligonuc1
6	604	19.0	1026	6 ABO22411	Abq22411 Oligonuc1
7	515	16.2	515	12 ACH73791	Ach73791 Human gen
8	514	16.1	514	6 ABT07220	Abt07220 Human Cpg
9	469.6	14.7	1026	6 ABO22412	Abq22412 Oligonuc1
10	469.6	14.7	1026	6 ABO22413	Abq22413 Oligonuc1
11	419	13.2	457	9 ACH15279	Ach15279 Human adu
12	384.6	12.1	388	3 AAC00887	Aac00887 Human sec
13	374.6	11.8	2494	3 AAZ46477	Aaz46477 Human T-b
14	374.6	11.4	3465	12 ADP28768	Adp28768 Human sec
15	363	11.4	389	2 AAQ39695	Aaq39695 Expressed
16	363	11.2	2425	11 AAQ59107	Aaq59107 Human bra
17	357	11.0	2425	11 ADL22644	Adl22644 Human dis
18	350.8	11.0	2883	3 AAZ90784	Aaz90784 Mouse ner
19	346	10.9	2061	3 AAZ90785	Aaz90785 Human ner

20	346	10.9	2061	12 ADQ19927	Adq19927 Human sof
21	342.8	10.8	3137	12 ADQ24019	Adq24019 Human sof
22	341.2	10.7	2756	14 ADY20406	Ady20406 DNA encod
23	341.2	10.7	2756	14 ADY19051	Ady19051 DNA encod
24	321	10.1	463	13 ACF88226	Acf88226 Human SIR
25	316.2	9.9	1608	4 AAC91988	Aac91988 Human T-b
26	316.2	9.9	1608	9 ADA02650	Ada02650 Human T-b
27	316.2	9.9	1608	9 ACC84950	Acc84950 Human T-b
28	316.2	9.9	1608	10 ADE95898	Ade95898 Human Tbx
29	316.2	9.9	1608	10 ADB72388	Adb72388 Human Tbx
30	316.2	9.9	1608	14 ADV43485	Adv43485 Human psy
31	316.2	9.9	2450	14 ADZ39787	Adz39787 Human T-b
32	316.2	9.9	2589	14 ADZ39787	Adz39787 Human T-b
33	316.2	9.9	2589	9 ADA02649	Ada02649 Human Tbx
34	316.2	9.9	2589	10 ADB72387	Adb72387 Human Tbx
35	316.2	9.9	2589	10 ADE95897	Ade95897 Human Tbx
36	316.2	9.9	2589	13 ADP55341	Adp55341 Human PRO
37	316.2	9.9	2589	14 ADX01430	Adx01430 Human tol
38	316.2	9.9	2589	14 ADY17747	Ady17747 DNA encod
39	316.2	9.9	2590	14 ADZ39814	Adz39814 Human T-b
40	298.6	9.4	1593	9 ADA02647	Ada02647 Mouse Tbx
41	298.6	9.4	1593	10 ADB72385	Adb72385 Mouse Tbx
42	298.6	9.4	1593	10 ADE95895	Ade95895 Mouse Tbx
43	298.6	9.4	2482	9 ADA02646	Ada02646 Mouse Tbx
44	298.6	9.4	2482	10 ADB72384	Adb72384 Mouse Tbx
45	298.6	9.4	2482	10 ADE95894	Ade95894 Mouse DNA

ALIGNMENTS

RESULT 1
ADL97817
ID ADL97817 standard, DNA, 3186 BP.
XX
XX ADL97817;
XX
DT 03-JUN-2004 (first entry)
XX
XX Human TBR1 DNA SeqID 3.
DE
XX human; gene; ds; neuropsychiatric disorder; bipolar disorder;
XX schizoprenia; alpha-type II calcium/calmodulin dependent protein kinase;
XX CMK1T-alpha; TBR1; Brachyury; neuroleptic; antidepressant; gene therapy;
KW mental disorder; mood disorder; psychosis; major depression.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 303..2351
FT /ftag= a
FT /product= "TBR1 protein"
XX
XX W02004020455-AZ.
XX
XX 11-MAR-2004.
XX
XX 26-AUG-2003; 2003WO-US026873.
XX
XX 28-AUG-2002; 2002US-0406879P.
XX
XX 27-FEB-2003; 2003US-0451306P.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Bunney WR, Jones EG, Molnar M;
PI WPI; 2004-239152/22.
XX
XX P-PSDB; ADL97818.
XX
XX Determining whether a subject has or is predisposed to a mental disorder,
XX comprises detecting the level of reagent that selectively associates with
XX a specified polynucleotide or polypeptide in a sample from the subject.
XX

PS Claim 1; SEQ ID NO 3; 85bp; English.

CC This invention relates to a novel method to determine whether a subject
CC has or is predisposed to a mental or neuropsychiatric disorder.
CC Specifically, it refers to identifying modulators of two genes expressed
CC in the central nervous system of patients suffering from bipolar disorder
CC or schizophrenia, namely alpha-type II calcium/calmodulin dependent
CC protein kinase (CAMKII-alpha) and TRB1 (a putative transcription factor
CC related to the Brachyury gene). The present invention describes screening
CC assays used to identify gene expression modulators, as well as
CC appropriate antibodies, agonists and antagonists thereof. Accordingly,
CC these compositions, which exhibit neuroleptic and antidepressant
CC activities can also be used for gene therapy purposes to treat the
CC aforementioned conditions, and mental disorders including mood disorders,
CC psychosis and major depression. This polynucleotide sequence is the human
CC TRB1 DNA of the invention.

CC XX Sequence 3186 BP; 692 A; 998 C; 774 G; 722 T; 0 U; 0 Other;

Query Match 100.0%; Score 3186; DB 12; Length 3186;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 3186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTGATTATCTTAATTATGCTATCTAATTAATTACTGTCAAGCTTAACCAATGG 60
DB 1 CAGGTGATTATCTTAATTATGCTATCTAATTAATTACTGTCAAGCTTAACCAATGG 60
QY 61 CAGGAGCCGTTTCATCGCGCTGCACAAAGCAAGATCAAAAGTAGCCCTTTCTGATTGC 120
DB 61 CAGGAGCCGTTTCATCGCGCTGCACAAAGCAAGATCAAAAGTAGCCCTTTCTGATTGC 120
QY 121 TGCATAGTGCATATGGCCCATCTCTTCCCAAGGAAAAAAGTAATCAAACTT 180
DB 121 TGCATAGTGCATATGGCCCATCTCTTCCCAAGGAAAAAAGTAATCAAACTT 180
QY 181 TGAGAAGCATTTGCTGGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
DB 181 TGAGAAGCATTTGCTGGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
QY 241 TTTATGATTAATAGCATTTAAAAACAGAGGACGGGAGCGCATGTTTCTAGAG 300
DB 241 TTTATGATTAATAGCATTTAAAAACAGAGGACGGGAGCGCATGTTTCTAGAG 300
QY 301 CTATGAGCTGAGGACGCTGCTTCTCTATCATGCTCTCAAGAAATTTCTCAATG 360
DB 301 CTATGAGCTGAGGACGCTGCTTCTCTATCATGCTCTCTCAAGAAATTTCTCAATG 360
QY 361 TGAGCAGCAGCTTACCCACATTCAGGCGGATCCGAGCTTGTCTTGCACGATCCCATTA 420
DB 361 TGAGCAGCAGCTTACCCACATTCAGGCGGATCCGAGCTTGTCTTGCACGATCCCATTA 420
QY 421 TCTCCAGCAGCTGACCAACTGAGAGAAATTCACCTTTGAAAAAATTCACAGGGGATGA 480
DB 421 TCTCCAGCAGCTGACCAACTGAGAGAAATTCACCTTTGAAAAAATTCACAGGGGATGA 480
QY 481 CGAATAGCAGTAAGCAATTTTCTGATCTCAAGAGCTCACAGGGGAGCGTCAGA 540
DB 481 CGAATAGCAGTAAGCAATTTTCTGATCTCAAGAGCTCACAGGGGAGCGTCAGA 540
QY 541 GAAATGAATCTCTCTCTCTTGAAGCGGGTCTTGAAGCTTGTCAAGTTTCAATGCT 600
DB 541 GAAATGAATCTCTCTCTCTTGAAGCGGGTCTTGAAGCTTGTCAAGTTTCAATGCT 600
QY 601 CTGCTGCAATGCTTACCTCTCTCTCAATCCAGGCAAGTGTGGGCACTGCTC 660
DB 601 CTGCTGCAATGCTTACCTCTCTCTCAATCCAGGCAAGTGTGGGCACTGCTC 660
QY 661 CCAAGTGCATGTTCCGCTACCCGCGCAGGACCGGCGGACCCCGCTTCTCATG 720
DB 661 CCAAGTGCATGTTCCGCTACCCGCGCAGGACCGGCGGACCCCGCTTCTCATG 720
QY 721 GGAGCCCTAGCCGCTACATGCGCCACACCCGGTCTATCAACAGAGCTTCAACAGCC 780
DB 721 GGAGCCCTAGCCGCTACATGCGCCACACCCGGTCTATCAACAGAGCTTCAACAGCC 780

DB 721 GGAGCCCTAGCCGCTACATGCGCCACACCCGGTCTATCAACAGAGCTTCAACAGCC 780
QY 781 TCTGTCCAACTCTCCGCGAGGAGATACCCCAAGGCGGCTACCCCTACACGAGAGT 840
DB 781 TCTGTCCAACTCTCCGCGAGGAGATACCCCAAGGCGGCTACCCCTACACGAGAGT 840
QY 841 ACAGCAGCTCTTACCAAGAGAGCTCCGTTTACCAAGTCTCTCTCCACCAAGCGGAGCTG 900
DB 841 ACAGCAGCTCTTACCAAGAGAGCTCCGTTTACCAAGTCTCTCTCCACCAAGCGGAGCTG 900
QY 901 TGCCCGCAAGCAACAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGCGC 960
DB 901 TGCCCGCAAGCAACAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGCGC 960
QY 961 ACCAAGGAGATGATCATCAACCAAGGAGGCGGATTTCTTTTAAGTTTA 1020
DB 961 ACCAAGGAGATGATCATCAACCAAGGAGGCGGATTTCTTTTAAGTTTA 1020
QY 1021 ACATTTCTGTCTGATCCACAGGCTCATTAATATTTTGTGATGTGATTTGGCGG 1080
DB 1021 ACATTTCTGTCTGATCCACAGGCTCATTAATATTTTGTGATGTGATTTGGCGG 1080
QY 1081 ATCCCAATCACTGAGGTTTCAAGAGGCAATGAGTTCTTGGCGCAAGCGACCA 1140
DB 1081 ATCCCAATCACTGAGGTTTCAAGAGGCAATGAGTTCTTGGCGCAAGCGACCA 1140
QY 1141 ATGTGCAAGGAATCGGGTCTATATGATCGGATTCGCCCAACACTGGGGCTGACTGGA 1200
DB 1141 ATGTGCAAGGAATCGGGTCTATATGATCGGATTCGCCCAACACTGGGGCTGACTGGA 1200
QY 1201 TGCGCAAGAAATCTCTTTGAAAAATTTAACTTACGACCAAGAGAGCTTCAATA 1260
DB 1201 TGCGCAAGAAATCTCTTTGAAAAATTTAACTTACGACCAAGAGAGCTTCAATA 1260
QY 1261 ACAATGGCAGATGGTGTATTAAGTCTTGGCAACATGACAGCCCGCTGATGAG 1320
DB 1261 ACAATGGCAGATGGTGTATTAAGTCTTGGCAACATGACAGCCCGCTGATGAG 1320
QY 1321 TGGAAGTGAAGAGGAGGCGGACGAGAGCACTAGCCAGCCGCGCGGTGAGAGCTTCA 1380
DB 1321 TGGAAGTGAAGAGGAGGCGGACGAGAGCACTAGCCAGCCGCGCGGTGAGAGCTTCA 1380
QY 1381 CTCTTCCCTGAGACTCAGTTCATCGCCGCTACCGGCTACCAAGACAGGATTTACACAC 1440
DB 1381 CTCTTCCCTGAGACTCAGTTCATCGCCGCTACCGGCTACCAAGACAGGATTTACACAC 1440
QY 1441 TGAATAATGATCACAACTTTTGCAAAAGATTCGGGATTAATATGACAGATCTTACA 1500
DB 1441 TGAATAATGATCACAACTTTTGCAAAAGATTCGGGATTAATATGACAGATCTTACA 1500
QY 1501 CCGGCTGTGACATGAGCCGCGTGAACCCCTGCGCCCAAGCATCGCGCGCTCGAGATCG 1560
DB 1501 CCGGCTGTGACATGAGCCGCGTGAACCCCTGCGCCCAAGCATCGCGCGCTCGAGATCG 1560
QY 1561 TGCCCGGGCCCGCTTACCGCATGCGCGGCTCTTCTCTGCAAGACCAAGTTCTGAGCACT 1620
DB 1561 TGCCCGGGCCCGCTTACCGCATGCGCGGCTCTTCTCTGCAAGACCAAGTTCTGAGCACT 1620
QY 1621 ACGCCAAGCCCGCTTACCCGCGCGCGGCGGCGGCTCCGCGGCTGACGACCGCA 1680
DB 1621 ACGCCAAGCCCGCTTACCCGCGCGCGGCGGCGGCTCCGCGGCTGACGACCGCA 1680
QY 1681 GCGTGCAGCAGCAACAGGGGTGCTGTGCGCGCAGACAGGCGGAGGACCCGCGCGGCT 1740
DB 1681 GCGTGCAGCAGCAGCAACAGGGGTGCTGTGCGCGCAGACAGGCGGAGGACCCGCGCGGCT 1740
QY 1741 CGCGCAACGCTGTGTTGACGCGCGCAACCAACCGGCTGAGCTTCTGCGGCT 1800
DB 1741 CGCGCAACGCTGTGTTGACGCGCGCAACCAACCGGCTGAGCTTCTGCGGCT 1800
QY 1801 ATGACAGGCGCAAGGACTTTCGCGGCAAGCGGCGCAAGCTGCTCTTACGCGGCGG 1860
DB 1801 ATGACAGGCGCAAGGACTTTCGCGGCAAGCGGCGCAAGCTGCTCTTACGCGGCGG 1860

[illegible]

Qy	2941	ACAAACCAGTCAATGCTTAACCTTTTTCCTTCCCTTGCTTTCTTCTGC	3000
Db	2941	ACAAACCAGTCAATGCTTAACCTTTTTCCTTCCCTTGCTTTCTTCTGC	3000
Qy	3001	TCTCATACTTTCTCTCTCTCTTTTAATTTCTGTGAGATAATTTCAAGGCTC	3060
Db	3001	TCTCATACTTTCTCTCTCTCTTTTAATTTCTGTGAGATAATTTCAAGGCTC	3060
Qy	3061	TAGAACAATGAATACTCAATGAGATGGGGTTCCACTTCTCCTCAATCGGTGATGA	3120
Db	3061	TAGAACAATGAATACTCAATGAGATGGGGTTCCACTTCTCCTCAATCGGTGATGA	3120
Qy	3121	AATATTACTAATGTGCCCTAATGACACAATAATAGTAGAGAAATCCAACCAACCTT	3180
Db	3121	AATATTACTAATGTGCCCTAATGACACAATAATAGTAGAGAAATCCAACCAACCTT	3180
Qy	3181	TAAAGG 3186	
Db	3181	TAAAGG 3186	

RESULT 2
AEA79130
ID AEA79130 standard; cDNA; 1793 BP.

DT 11-AUG-2005 (first entry)

Human apoptosis-associated cDNA SEQ ID 378.

KM apoptosis; cell death; cytoskeletal; neuroprotective; immunosuppressive;
KM antineoplastic; antiarthritic; dermatological; antiinflammatory;
KM hepatotropic; virucide; nootropic; anticonvulsant; antiParkinsonian;
KM vasotropic; cerebroprotective; antialcoholic; gene therapy; tumour;
KM autoimmune disease; degenerative disease; rheumatoid arthritis; diabetes;
KM carcinosarcoma; sarcoma; multiple sclerosis; viral infection; leukaemia;
KM lupus; hepatitis; influenza viruses; Alzheimer's disease;
KM Huntington's disease; Parkinson's diseases; reperfusion injury; stroke;
KM alcoholic liver disease; human; gene; ss.

Homo sapiens.

WO2003058021-A2.

17-JUL-2003.

13-JAN-2003; 2003WO-EP000270.

11-JAN-2002; 2002DE-01000856.

(XANT-) XANTOS BIOMEDICINE AG.

Koenig-Hoffman K, Kazinski M, Schaefer R, Kesper B;

WPI; 2003-542134/51.

New nucleic acids involved in apoptosis, useful for diagnosis and treatment of e.g. tumors and degenerative disease, also related proteins, antibodies and modulators.

Claim 1a; SEQ ID NO 378; 517pp; German.

This invention describes novel nucleic acid molecules that are associated with apoptosis and encode a polypeptide and are derived from a normalised gene library (emryonic or liver) or clone collections, and the extent of apoptosis measured by cell death detection assay or the CPG assay (measuring loss of membrane integrity). The products of the invention have cytostatic, neuroprotective, immunosuppressive, antineumatic, antiarthritic, dermatological, antiinflammatory, hepatotropic, virulide, neurotropic, anticonvulsant, antiparkinsonian, vasotropic, cerebroprotective and antialcoholic activity and can be used for gene therapy. The polynucleotides also related vectors, hosts (or their

CC extract(s), encoded polypeptide (or their receptors) and/or agents that
CC inhibit their activity (including antisense sequences) are used for
CC treatment or prevention of tumours, autoimmune or degenerative diseases
CC and viral infections, specifically leukaemia, carcinoma, sarcoma,
CC multiple sclerosis, rheumatoid arthritis, diabetes, lupus, or infection
CC with hepatitis or influenza viruses, Alzheimer's, Huntington's or
CC Parkinson's diseases, reperfusion injury, stroke and alcoholic liver
CC disease. Detection of the polymucleotides and derived polypeptides can
CC also be used for diagnosis of these diseases. This sequence encodes an
CC apoptosis-associated protein described in the disclosure of the
CC invention.
CC
CC
XX
XX

Sequence 1793 BP; 385 A; 586 C; 471 G; 351 T; 0 U; 0 Other;

Query Match 56.3%; Score 1793; DB 11; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 225 GTCCTGGATTCTCTGTTTATGATAAATAGACCTTTAAAAACAGGAGCGGAGCGAG 284
DB 1 GTCCTGGATTCTCTGTTTATGATAAATAGACCTTTAAAAACAGGAGCGGAGCGAG 60
OY 285 TGTTCAGGTTCTAGAGCTATGAGCACTGAGCACTGCTTCTCTTATCATGCTCTCC 344
DB 61 TGTTCAGGTTCTAGAGCTATGAGCACTGAGCACTGCTTCTCTTATCATGCTCTCC 120
OY 345 AAGAAATTTCTCAATGTGAGCAGCACTCACTTCAAGCGGATCCGAGCTTGTCTTG 404
DB 121 AAGAAATTTCTCAATGTGAGCAGCACTCACTTCAAGCGGATCCGAGCTTGTCTTG 180
OY 405 CACGATCATCCCATATCTCCGACCACTGACCACTGAGAGAGATTCACCTTTGAAAAA 464
DB 181 CACGATCATCCCATATCTCCGACCACTGAGAGAGATTCACCTTTGAAAAA 240
OY 465 ATTACCAAGGGAGTACGAAATGTCAGATACAGCAATTTTCTGACTCCAGAGCTCA 524
DB 241 ATTACCAAGGGAGTACGAAATGTCAGATACAGCAATTTTCTGACTCCAGAGCTCA 300
OY 525 CCAAGGGAGCTCCCAAGAAAGTAACTCTCTGCTTGTGAGGGGTCTTGAAGTTTGT 584
DB 301 CCAAGGGAGCTCCCAAGAAAGTAACTCTCTGCTTGTGAGGGGTCTTGAAGTTTGT 360
OY 585 CACAGTTTCAGATGCTGCTGCTGAGTGCCTCTCTCTCACTGACGAGCCAGCAG 644
DB 361 CACAGTTTCAGATGCTGCTGCTGAGTGCCTCTCTCTCACTGACGAGCCAGCAG 420
OY 645 TCTGGGACCACTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAGACCGGCGAC 704
DB 421 TCTGGGACCACTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAGACCGGCGAC 480
OY 705 CCGGCTTCTCTCAATGGGAGCCCTAGCCGCTACATGGCCCAACCCGGTCAATCCAC 764
DB 481 CCGGCTTCTCTCAATGGGAGCCCTAGCCGCTACATGGCCCAACCCGGTCAATCCAC 540
OY 765 GAGAGCTCAACAAGGCTCTGTCCTGCACTCTGCGGAGGATACCCGAGCGGCTAGC 824
DB 541 GAGAGCTCAACAAGGCTCTGTCCTGCACTCTGCGGAGGATACCCGAGCGGCTAGC 600
OY 825 CCTTACCCACAGACAGTACGGCCACTCTACCAAGAGTCCGTTTACAGTTCTCTCC 884
DB 601 CCTTACCCACAGACAGTACGGCCACTCTACCAAGAGTCCGTTTACAGTTCTCTCC 660
OY 885 ACCCAGCCGGGCTGTGCTCCGCGCAAGACAGGTGTACCTGTGCAACAGGCCCTTTGG 944
DB 661 ACCCAGCCGGGCTGTGCTCCGCGCAAGACAGGTGTACCTGTGCAACAGGCCCTTTGG 720
OY 945 CTGAATTTCAACCGGACCAAAAGGAGTGCATCAACCAACAGGAGGCGCATGTTT 1004
DB 721 CTGAATTTCAACCGGACCAAAAGGAGTGCATCAACCAACAGGAGGCGCATGTTT 780
OY 1005 CTTTTTTTAAAGTTTAAATTTCTGCTCGATCCCAAGGCTCATTAATATTTTGTG 1064
DB 781 CTTTTTTTAAAGTTTAAATTTCTGCTCGATCCCAAGGCTCATTAATATTTTGTG 840

OY 1065 GATGTGATTTTGGGGGATCCCAATCACTGAGGTTTGAAGAGCAATGGTCTCTGC 1124
DB 841 GATGTGATTTTGGGGGATCCCAATCACTGAGGTTTGAAGAGCAATGGTCTCTGC 900
OY 1125 GGCACACGGGACACCAATGTGCAAGAAATGCGGTCTATATGATCCGATTTCCCAAC 1184
DB 901 GGCACACGGGACACCAATGTGCAAGAAATGCGGTCTATATGATCCGATTTCCCAAC 960
OY 1185 ACTGGGGCTCATGTGATGCGCCAAAGAAATCTTTTGAATAATTAACCTTACGACAC 1244
DB 961 ACTGGGGCTCATGTGATGCGCCAAAGAAATCTTTTGAATAATTAACCTTACGACAC 1020
OY 1245 AAGAGGCTTCAATTAACATAGGCGAGATGGTGTATTACAGTCCCTGCAACATGACAG 1304
DB 1021 AAGAGGCTTCAATTAACATAGGCGAGATGGTGTATTACAGTCCCTGCAACATGACAG 1080
OY 1305 CCGGCTGATGTGTGGAAGTGAAGAGACGGACCGGACCGAGGACACTAGCCAGCCGCG 1364
DB 1081 CCGGCTGATGTGTGGAAGTGAAGAGACGGACCGGACCGGACCGAGGACACTAGCCAGCCGCG 1140
OY 1365 CGGTCGACAGCTTCACTTTCCTGAGACTCACTTCACTCCGCTGACCGCTTACAGAAC 1424
DB 1141 CGGTCGACAGCTTCACTTTCCTGAGACTCACTTCACTCCGCTGACCGCTTACAGAAC 1200
OY 1425 ACGGATTTACACAACTGAAATAGATCAACACCTTTTGAAGAAAGATTTCCGGATAT 1484
DB 1201 ACGGATTTACACAACTGAAATAGATCAACACCTTTTGAAGAAAGATTTCCGGATAT 1260
OY 1485 TATGACACGATCTACACCGGCTGTGATGAGACCGCTGACCCCTGAGCCCAAGACTCG 1544
DB 1261 TATGACACGATCTACACCGGCTGTGATGAGACCGCTGACCCCTGAGCCCAAGACTCG 1320
OY 1545 CCGGCTGACAGATGTCGCCGGGGCCCGCTAGCGCATGCGGCTCTTTCTGACAGAC 1604
DB 1321 CCGGCTGACAGATGTCGCCGGGGCCCGCTAGCGCATGCGGCTCTTTCTGACAGAC 1380
OY 1605 CAGTTTGTGAGCACTAGCGCAAGGCGGCTTCCACCGGGGCGGGGCGGGGCGGGG 1664
DB 1381 CAGTTTGTGAGCACTAGCGCAAGGCGGCTTCCACCGGGGCGGGGCGGGGCGGGG 1440
OY 1665 CCGGGTACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1724
DB 1441 CCGGGTACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1500
OY 1725 GACCGGCGGCGGCTTCCGCGCAACGCTGTTTGTGACCGCGGCAACACCGGCTGAC 1784
DB 1501 GACCGGCGGCGGCTTCCGCGCAACGCTGTTTGTGACCGCGGCAACACCGGCTGAC 1560
OY 1785 TTGGGGCTTCCGCTGATGACAGGGCCAGGACCTTGGCGGGCAACGGGGCGAGCTGCTC 1844
DB 1561 TTGGGGCTTCCGCTGATGACAGGGCCAGGACCTTGGCGGGCAACGGGGCGAGCTGCTC 1620
OY 1845 TCTTACCGGCGGCGGCGGAGGAGGCGCTGCGCTGACAGGCTGACAGCTGAGCGCGC 1904
DB 1621 TCTTACCGGCGGCGGCGGAGGAGGCGCTGCGCTGACAGGCTGACAGCTGAGCGCGC 1680
OY 1905 CCGCTGAGCTTACGCGCAACCGCTCGGGCTGAGGCGCGCGAGTCCCGGAGTACTGC 1964
DB 1681 CCGCTGAGCTTACGCGCAACCGCTCGGGCTGAGGCGCGCGAGTCCCGGAGTACTGC 1740
OY 1965 GGCACCAAGTCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2017
DB 1741 GGCACCAAGTCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1793

RESULT 3
ID ABA14582/c
ID ABA14582 standard; DNA; 6878 BP.
XX ABA14582;
XX
DT 23-JAN-2002 (first entry)

[illegible]

OY	1910	CGGCTACTACGCGCAGACCCCTGTGCGGCTGCGGCGGCCCGCAGATCTCCCGCAGATCTACGCGCAC	1969
Db	2087	CGGCTACTACGCGCCACCCGCTCGGCGTGGGGCGCCCGCAGTCTCCCGCAGATCTACGCGCAC	2028
OY	1970	CAAGTCGGGAGCTCGGTGTGCTGCTCTCTGAGCCAAACAGCGCGCGCGCCGCGCGCATAGCG	2029
Db	2027	CAAGTCGGGAGCTCGGTGTGCTGCTCTCTGAGCCAAACAGCGCGCGCGCGCCGCGCGCATAGCG	1968
OY	2030	CGGCGCCAAATCTCTACTCTGGGCGAGAGGCGCGAGGCGCTTGCGCGCGCGAGCGCTTGCGGCT	2089
Db	1967	CGGCGCCAAATCTCTCTACTCTGGGCGAGAGGCGCGAGGCGCGCTTGCGCGCGCGAGCGCTTGCGGCT	1908
OY	2090	GGCGCGCGCGCGCGCGCGCAGAGCGCAGAGCCCAAGACCTGTCCGATTCAGACTGATTCGA	2149
Db	1907	GGCGCGCGCGCGCGCGCGCAGAGCGCAGAGCCCAAGACCTGTGTGCGATTCAGACTGATTCGA	1848
OY	2150	GACGCGCTCTCGATCAATCAAGTTCATCGACTCTCCAGCGACTCGGGGATTTTACGAGCGGCCAA	2209
Db	1847	GACGCGCTCTCGATCAATCAAGTTCATCGACTCTCCAGCGACTCTGGGGGATTTTACGAGCGGCCAA	1788
OY	2210	GGCGAGCGCGATCTTCGCGCGCGCGCAGCGCCGATTCGAGATTCGTCTCCCGCTCAAGAG	2269
Db	1787	GGCGAGCGCGATCTTCGCGCGCGCGCAGCGCCGATTCGAGATTCGTCTCCCGCTCAAGAG	1728
OY	2270	CGAGGTCTCGGCGCCAGCGGAGCTTCGAGAAAGAACTGCGCAGAGCAATTAAGCGGCTACTA	2329
Db	1727	CGAGGTCTCGGCGCCAGCGGAGCTTCGAGAAAGAACTGCGCAGAGCAATTAAGCGGCTACTA	1668
OY	2330	TGCGTCTTACTTCGACACGACTAAGCGCGCCCTGTCCCGCCCGGCGCGCGCGCGGAC	2389
Db	1667	TGCGTCTTACTTCGACACGACTAAGCGCGCCCTGTCCCGCCCGGCGCGCGCGCGGAC	1608
OY	2390	CCGAGCAGAGCCCTTCACAGGCTCTTCCCGAGCTCGCGCTCCCGCACACTCTCTCTTGCGCAC	2449
Db	1607	CCGAGCAGAGCCCTTCACAGGCTCTTCCCGAGCTCGCGCTCCCGCACACTCTCTCTTGCGCAC	1548
OY	2450	CCACTCAATTTTATTTGAACCTCGATGGCCGCTCTGACAGCAATTAAGCGAGGCTTCGAGC	2509
Db	1547	CCACTCAATTTTATTTGAACCTCGATGGCCGCTCTGACAGCAATTAAGCGAGGCTTCGAGC	1488
OY	2510	GTGATTTTAAACCTTTTTTGGACACAGACATCTCTGCAATTAAGTCACCGACTTCAACTTGG	2569
Db	1487	GTGATTTTAAACCTTTTTTGGACACAGACATCTCTGCAATTAAGTCACCGACTTCAACTTGG	1428
OY	2570	CTGTAAACCTTTTGGTTTTCTTACTTACTCTCTCTCTGTGGAGTTATCTCTCAAAATTC	2629
Db	1427	CTGTAAACCTTTTGGTTTTCTTACTTACTCTCTCTCTGTGGAGTTATCTCTCAAAATTC	1368
OY	2630	CCCTTCCCGCTGTCTTCTCTTACTCTCTACTTCTCTTTCTGTGTAATGAACCTTTCAAC	2689
Db	1367	CCCTTCCCGCTGTCTTCTCTTACTCTCTACTTCTCTTTCTGTGTAATGAACCTTTCAAC	1308
OY	2690	TTTATGAGAGACTGTGGGAGTCTCTGTACGAGCAGACAGATTCCGACCCGCGCAAGTCTCGGCC	2749
Db	1307	TTTATGAGAGACTGTGGGAGTCTCTGTGTACGAGCAGACAGATTCCGACCCGCGCAAGTCTCGGCC	1248
OY	2750	TCGACATTTAACCATAGATGTTGACTCTTAGAAGCTTGACCCACCGACGCGTCTTTCTT	2809
Db	1247	TCGACATTTAACCATAGATGTTGACTCTTAGAAGCTTGACCCACCGACGCGTCTTTCTT	1188
OY	2810	ATCCCGCAGTGGATGAGATGCAATGCAATGAGATGGTATGTTAATTTTATTTAGTGAACA	2869
Db	1187	ATCCCGCAGTGGATGAGATGCAATGCAATGAGATGGTATGTTAATTTTATTTAGTGAACA	1128
OY	2870	AAGCCTGTGAATGATTGTACATAGTGTAAATTTATTTGTAAAGATGAGCTAGTTTTAT	2929
Db	1127	AAGCCTGTGAATGATTGTACATAGTGTAAATTTATTTGTAAAGATGAGCTAGTTTTAT	1068
OY	2930	CTGCTCAAGGCAAAAACAGTTATGTCTTAACTTTTTCTTTCTCTTCTTTGCTTT	2989
Db	1067	CTGCTCAAGGCAAAAACAGTTATGTCTTAACTTTTTCTTTCTCTTCTTTGCTTT	1008

|||||
DB 721 GAGGGGATCTCGCGCGGACACGCCCGGTGTCGAGAGTTGTGTCGCCGCTCAAGGGA 780
OY 2273 GGTGCTGGCCCCGCGGAGCTGCGAGAACTGCCCGCAAGACATTAACGGCTACTAAGG 2332
DB 781 GGTGCTGGCCCCGCGGAGCTGCGAGAACTGCGCAAGACATTAACGGCTACTAAGG 840
OY 2333 CTTTACTGCGACAGCTAGGCGCG 2356
DB 841 CTTTACTGCGACAGCTAGGCGCG 864
RESULT 5
ABO22410/c
ID ABO22410 standard; DNA; 1026 BP.
XX
AC ABO22410;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9001.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; de.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56bp + Sequence Listing; 56bp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABO22410-
CC ABO24121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1026 BP; 140 A; 128 C; 429 G; 329 T; 0 U; 0 Other;

Query Match 19.0%; Score 604; DB 6; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9.1e-114;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;
OY 1490 CACGATCTACACCGGCTGTGACATGAGACCGCTCAACCCCTTCGCGCAACGACTGCGCGG 1549
DB 949 CAAATCTACACCGACATTAACATTAACCGCTCAACCCCTTCGCGCAACGACTGCGCGG 890
OY 1550 CTGCGAGATGCTGTCGCCGGGGCCGCTACGACATGAGCGGCTCTTCTTCGACAGACGATT 1609
DB 889 CTGCGAANTGCTACCGCAAAACCGCTACGACATTAACGACTCTTCTTCAAAACCAATT 830
OY 1610 CGTAGCACTACGCGCAAGGCGCGCTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGG 1669
DB 829 CGTAAACAACTACCGCAAAACCGCTTCAACCGCAAGC--GACGCGAACCGGAACCGAA 772
OY 1670 TACGAGCCGACGAGCTGCGGCAACCAACGAGCTGCTGTCGCGGACGCGGCGCGG 1729
DB 771 TACGAGCCGCAACGTACCGGCAACCAACGAACTACTATGCGCGCAACCAACCGAAACCC 712
OY 1730 GGGGCGCGCGCGCGCGCAACGCTGTTGTGACGCGCGGCAACCAACCGGCTGACTTCG 1789
DB 711 GAACGCG--CTTCGCGCGACGCTAATTATTAACGCGCAACCAACCGACTTAACCTTCG 653
OY 1790 GGCCTCGGCTATGACACGCGCAACGCACTTCGCGGCAACGCGGCAACGCTGCTCTTTA 1849
DB 652 GACCTCGACCTATTAACACGACACGAACTTCGCGCAACGCGGCAACGCTACTCTCTTA 593
OY 1850 CGCGGCGCGGCGCTGAAAGCGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCGCT 1909
DB 592 CGCGACGACGAACTTAACAAAGCT--ACGCTACAACTACAACTACAACTACAACTAC 534
OY 1910 CGGCTACTACGCGCAACCGCTGCGGCGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGG 1969
DB 533 GCACTACTACGCGCAACCGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
OY 1970 CAAATCGGAGCTGAGTGTCTGCTGCTGCTGCGCGCAACGCGCGCGCGCGCGCATGCG 2029
DB 473 CAAATCGAAGCTGAGTGTCTGCTGCTGCTGCGCGCAACGCGCGCGCGCGCGCATGAC 415
OY 2030 CGGCGCCCAATCTCTAATCTGAGGAGGAGGAGGCGCTGCGCGCGCGCGCGCGCGCT 2089
DB 414 GACGCGCAATCTCTAATCTGAGGAGGAGGAGGAGGCGCTGCGCGCGCGCGCGCGCGCT 355
OY 2090 GCGCGCGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2149
DB 354 ACCGCGC--GCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 296
OY 2150 GACGCGCTCTGATCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2209
DB 295 AACGCGCTCTGATCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 236
OY 2210 GCGGAGCGGATCTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2269
DB 235 ACGAAACGAAATCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 176
OY 2270 CGAGTGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2329
DB 175 GAAATATTAACCGCAACCAATTAAGAAACCAATTAAGAAACCAATTAAGAAACCAATTA 116
OY 2330 TGGCTTACTCTGACAGCTAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGGAC 2389
DB 115 TAACTTCTACTGCAACATTAACGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGGAC 56
OY 2390 CCGACGCGCGCGCTGACAGCTTTCGCCGCTGCGCGCTGCGCGCACTCTCTT 2443
DB 55 CCGAACCAACCGCTGACAGCTTTCGCCGCTGCGCGCTGCGCGCACTCTCTT 2
RESULT 6
ABO22411
ID ABO22411 standard; DNA; 1026 BP.
XX

AC ABQ22411;
 XX 12-JUL-2002 (first entry)
 DT
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 9002.
 DE
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 PN W0200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP010074.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 PR 05-SEP-2000; 2000DE-01044543.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.
 XX
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridization to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridized to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ1410-
 CC ABQ4121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 CC
 SQ Sequence 1026 BP; 329 A; 429 C; 128 G; 140 T; 0 U; 0 Other;
 XX
 XX
 Query Match 19.0%; Score 604; DB 6; Length 1026;
 Best Local Similarity 80.5%; Pred. No. 9.1e-114;
 Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;
 XX
 QY 1490 CACGATCTACACCGCTGTGACATGAGCCGCTGACCCCTCGCCCAAGACTCGCGCG 1549
 DB 78 CAAATCTACACCGCTATTAATGATTAACCGCTTAACCCCTCGCCCAAGACTCGCGCG 137
 QY 1550 CTGCGAGATGTGCTGCGGCGCGCTGACGCGCGCTCTTTCTTGAAGACCAAT 1609
 DB 138 CTGCGAAATCTGTAACCGGAAACCGCTACGCGCAATGACTTTCTTCAAAACCAATT 197
 QY 1610 CGTGAACAATGACGCAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCG 1669
 DB 198 CGTAAACAATGACGCAAGCGCGCTTCCACCGGAAACCGCGCGCGCGCGCGCGCG 255

QY 1670 TACGGAACGAGCGTGCGCCGACACCAACGGGCTGCTGTGCGCGGACGCGGACCC 1729
 DB 256 TACGGAACGAGCGTGCGCCGACACCAACGGGCTGCTGTGCGCGGACGCGGACCC 315
 QY 1730 GGGGCGCGCGCTGCGCGGACGCTGCTGTGAGCGCGCGGACCAACCGGCTGCACTTCC 1789
 DB 316 GAAAGCGC-CTGCGCGGACGCTAATTATTAACGCGGACCAACCGGCTGCACTTCC 374
 QY 1790 GGGCTGCGCTTATTAACGCGGACGCTTCCGCGGGAAGCGCGGACGCGGCTGCTTA 1849
 DB 375 GACTTCGACCTTATTAACGAGACGAACTTCCGAAACGCGGACGCGGCTGCTT 434
 QY 1850 CGGCGCGCGCGGCTGAAAGCGCTGCGCTGCAAGGCTGCAAGGCTGCAAGGCTGCGCGCT 1909
 DB 435 CGGACGAGCGAAGCTTAAACGCT-AGGCTACAACTTACAACTTACCGCGCGCT 493
 QY 1910 CGGCTACTACGCGGACGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1969
 DB 494 GCACTACTAGCGCGGACCGCTGCACTTAAACGCGCGGCAATCCCGCAATCTACGACAC 553
 QY 1970 CAAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029
 DB 554 CAAATCGAACTGCACTGCACT-CCCTACTAACCGCAACGCGCGGACGCGCGGCTTAAAC 612
 QY 2030 CGGCGCAATCCCTACTGCGGCGGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCT 2089
 DB 613 CGAGCGCAATCCCTACTGCGGCGGAGAGCGGCGCGCGCGCGCGCGCGCGCGCT 672
 QY 2090 GCGCGCGCGCGCGCGCGGAGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGG 2149
 DB 673 ACCGCGCGCGCGGAGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGG 731
 QY 2150 GACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2209
 DB 732 AACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
 QY 2210 GCGGAGCGGAGTTCGCGCGCGGACGAGCGCGGCTGCGGAGTTCGCGCGGAGTTCGCG 2269
 DB 792 AGGAAAGAGTTCGCGCGCGGACGAGCGCGGCTGCGGAGTTCGCGCGGAGTTCGCG 851
 QY 2270 CGAGGTGCTGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 2329
 DB 852 CGAATATTAACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 911
 QY 2330 TGGCTTACTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 2389
 DB 912 TAACTTACTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 971
 QY 2390 CCCAGCGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 2443
 DB 972 CCAGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1025
 XX
 XX
 RESULT 7
 ACH73791
 ID ACH73791 standard; DNA; 515 BP.
 XX
 AC ACH73791;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #6986.
 DE
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00029386.

[illegible]

Db	181	CCAGCTCCGCCTCCCCACACTCTCTTGGCGACCACTCATTTATTGACCCCTGGAG	240
Qy	2476	GGCGTCTGACGGAATTAAGTCAGAGGTCTCCGAGCGTGATTTTAACCTTTTTCACAGCA	253
Db	241	GCCGCTCGACGGAATTAAGTCAGAGGTCTCCGAGCGTGATTTTAACCTTTTTCACAGCA	300
Qy	2536	GTCCTCTGGAATTAAGTCACCGACCTTCAACTTTGCTGTAAACCTTTGGTTTTCGACTT	259
Db	301	GTCTCTGGAATTAAGTCACCGACCTTCAACTTTGCTGTAAACCTTTGGTTTTCGACTT	360
Qy	2596	ACTCTTCTCTGTGGAGTTATCTCTCAATTTCCCTCCCTCGTCTTCTTCACTT	265
Db	361	ACTCTTCTCTGTGGAGTTATCTCTCAATTTCCCTCCCTCGTCTTCTTCACTT	420
Qy	2656	CCTACTTCTCTTCTTGTGAATGAACCTCTTCACTTTAGAGACCTGGGCACTCTGTCA	271
Db	421	CCTACTTCTCTTCTTGTGAATGAACCTCTTCACTTTAGAGACCTGGGCACTCTGTCA	480
Qy	2716	GGCAGCAGCATTCGACCCGCGAGTCTCGGCGT	2750
Db	481	GGCAGCAGCATTCGACCCGCGAGTCTCGGCGT	515
RESULT 8			
ABT07220/C	standard; DNA, 514 BP.		
AC	ABT07220;		
XX	ABT07220;		
XX	07-NOV-2002 (first entry)		
DE	Human Cpg-island rich DNA sequence SEQ ID No 37.		
XX	Cpg island; methylated; methylation-sensitive restriction enzyme; cancer;		
KW	malignant cell; tumour; human; ds.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200260318-A2.		
XX	08-AUG-2002.		
PD	31-JAN-2002; 2002WO-US003077.		
XX	31-JAN-2001; 2001US-007753398.		
XX	(OHIS) UNIV OHIO STATE RES FOUND.		
PA	Plaas C;		
XX	WPI; 2002-627436/67.		
DR	Identifying methylated Cpg islands in malignant cells, useful for		
XX	diagnosing or classifying cancer, comprises the use of infrequent		
PT	cleaving, methylation-sensitive restriction enzymes and gel		
PT	electrophoresis.		
XX	Claim 17; Page 86; 13pp; English.		
PS	The invention relates to a method for identifying Cpg islands which are		
XX	preferentially methylated in malignant cells, which comprises obtaining		
CC	genomic DNA from malignant and non-malignant cells, digesting this DNA		
CC	with infrequently-cutting, methylation-sensitive restriction enzymes to		
CC	form restriction fragments, electrophoresing these fragments and		
CC	comparing the intensity of these patterns to control cell restriction		
CC	fragments. The method is useful in diagnosing cancer, in classifying		
CC	tumours, in providing prognoses for cancer patients, and in identifying		
CC	new DNA clones within a library based on the methylation status of Cpg		
CC	dinucleotides. The isolated polynucleotides and oligonucleotides can also		
CC	be used as probes to identify genes whose expression is increased or		
CC	decreased in cancerous tissues. This polynucleotide sequence represents a		
CC	Cpg-island rich DNA sequence relating to the invention		

XX Sequence 514 BP; 66 A; 175 C; 207 G; 66 T; 0 U; 0 Other;
SQ
Query Match 16.1%; Score 514; DB 6; Length 514;
Best Local Similarity 100.0%; Pred. No. 2.1e-95;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GCTGTGACATGGAACCGCTGACCCCTGCGCCCAACGACTGCGCGCTGACGATGTC 1563
DB 514 GCTGTGACATGGAACCGCTGACCCCTGCGCCCAACGACTGCGCGCTGACGATGTC 455
QY 1564 CCGGGGCGCGCTACGCGCATGCGCGCTCTTCTGTGAGAACAGTTGAGCACTACG 1623
DB 454 CCGGGGCGCGCTACGCGCATGCGCGCTCTTCTGTGAGAACAGTTGAGCACTACG 395
QY 1624 CCAAGGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683
DB 394 CCAAGGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
QY 1684 TGCCTGACACCAACGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1743
DB 334 TGCCTGACACCAACGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
QY 1744 CGCAAGCTGTGTTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1803
DB 274 CGCAAGCTGTGTTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 215
QY 1804 ACACGCGCACGGAATTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1863
DB 214 ACACGCGCACGGAATTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 155
QY 1864 TGAAGGCGCTGCGCGCTGACAGGCTGACGCTGACGCTGCTGCTGCTGCTGCTGCTG 1923
DB 154 TGAAGGCGCTGCGCGCTGACAGGCTGACGCTGACGCTGCTGCTGCTGCTGCTGCTG 95
QY 1924 ACCCGTGGGCTGGGGCG 1983
DB 94 ACCCGTGGGCTGGGGCG 35
QY 1984 TGTGCTGCTGTGCGCGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2017
DB 34 TGTGCTGCTGTGCGCGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1

RESULT 9
ABQ22413/c
ID ABQ22413 standard; DNA; 1026 BP.
AC ABQ22413;
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9004.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001MO-BP010074.
XX
XX 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
XX (BPIG-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K, Gueutig D;

XX WPI; 2002-371829/40.
DR
XX
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridized to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridization to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridized to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ4121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1026 BP; 441 A; 311 C; 128 G; 146 T; 0 U; 0 Other;
Query Match 14.7%; Score 469.6; DB 6; Length 1026;
Best Local Similarity 71.7%; Pred. No. 3.2e-86;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

QY 1491 ACGATCTACGCGGCGTGAATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1550
DB 948 ACGATCTACGCGGCGTGAATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
QY 1551 TCGCAGATGTCGCCG 1610
DB 888 TCGTATGATGTCGTCG 829
QY 1611 GTAGCACTACGCAAGCG 1670
DB 828 GTAGCACTACGCAAGCG 771
QY 1671 ACGACCGCGAGGTCGCCG 1730
DB 770 ACGACCGCGAGGTCGCCG 711
QY 1731 GCG 1790
DB 710 GCG 652
QY 1791 GCGTGGCTGATGACG 1850
DB 651 GCGTGGCTGATGATGACG 592
QY 1851 GCGGCG 1910
DB 591 GCGGCG 533
QY 1911 GCGTACTACG 1970
DB 532 GCGTACTACG 473
QY 1971 AAGTCGGGCTGCGTGTGCG 2030
DB 472 AAGTCGGGCTGCGGTG-TTTTGTGTTGTTTATAGGTCGCGCGCGCGCGCGCGCGCGCG 414
QY 2031 GCGCGCAATCCCTAAGTGGCGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2090

```
Db 413 GCGCTTATTTTATTTGGCGAGAGGTCCAGGGTTTGTCGTGACGTTTCGTGCTG 354
Qy 2091 CCGCCGCGCGCGCGCGAGAGCGCCAAAGCACTGTCGATTCAGCTGATCGAG 2150
Db 353 TCGTTC-CCGTCGTGAGAGACGTAAAGTTTGAAGTTTGAATTTAGTTGATCGAG 295
Qy 2151 ACGCCCTCTCGATCAAGTCATCGACTCCAGCGACTGGGGATTAGAGAGCGCAAG 2210
Db 294 ACGTTTATTTTTCATTAAAGTTTATCAATTTTACGATTCGGGGATTAGAGATTAG 235
Qy 2211 CGAGCGGAGTCTCCGCGCGCCAGACGCGCGTGTCCAGAGATTCTCCCGCTCAAGAGC 2270
Db 234 CGGAGCGGATTTCTCGGTCGATACGTTTCGTGTCGAGAGTTCTTTTCGTTAAAGC 175
Qy 2271 GAGGTCTGCGCCAGCGGAGCTGCGAGAAAGAACTGCGCCAGAGCAATTAGCGGCTACTAT 2330
Db 174 GAGGTGTTGTTTACCGGAGATTGCGAAGAAATTCGTTAAAGATATTAGCGGTTATTAT 115
Qy 2331 GCGTCTACTCGCAGAGTACGCGCCGCTCGCGCGCGCGCGCGCGCGCGCGAGCC 2390
Db 114 GGTTTTATTCGTAAAGTTAGGTCGTTTGTTCGTTCGTTCGTTTCGTTCGAGATT 55
Qy 2391 CCAGCCAGCGCCCTCAGAGCTCTTCCCAAGCTCCGCTCCCACTCTCTCTT 2444
Db 54 TTAGTTAGTTTATTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTT 1
```

```
RESULT 10
ABQ22412
ID ABQ22412 standard; DNA; 1026 BP.
XX
AC ABQ22412;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9003.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
FN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PE 01-SEP-2001; 2001WO-BP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
```

degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABQ3410-CC ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention

Seq Sequence 1026 BP; 146 A; 128 C; 311 G; 441 T; 0 U; 0 Other;

Query Match 14.7%; Score 469.6; DB 6; Length 1026;
Best Local Similarity 71.7%; Pred. No. 3.2e-86;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

```
Qy 1491 ACGATCTACACCGGCTGTGACATGACCCGCTGACCCCTGCCCAAGACTCGCGCGC 1550
Db 79 AGGATTATATATCGGTTGTGATATGATCGTTGATTTTCGTTTACGATTCTGCGT 138
Qy 1551 TCGAGATCGTACCGCGCGCGCGCTACGCGCATGCGCGCTCTTCTGACAGACCAATT 1610
Db 139 TCGTAGATCGTTCGCGGCTTCGTTACGTTACGTTACGTTTCGTTTGTAGGATTAGTTC 198
Qy 1611 GTGAGCAACTAGCCCAAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGT 1670
Db 199 GTGAGTAAATTAAGTTAAGTTAGTTTATTCGCGCGCGCG--CGCGGATTCGCGT 256
Qy 1671 ACGAGCGGAGCGGCGCGGCGGACACCAACGCGGCTGCTGTGCGCGGAGCGGAGACCG 1730
Db 257 ACGAGTCGACGCGTGTGATTAACGCGTGTGTCGTGTGATGATGATGATGATGATTCG 316
Qy 1731 GCGCGCGCTGCGCGCAACGCTGTTGTGACGCGCGGCAACACGCGTGAATTCGCG 1790
Db 317 GCGCGG--TTTGTGTGTAACGTTGTTGTGACGTCGTTAATATCGTTGATTTGCGG 375
Qy 1791 GCTTCGCTATGACAGCGGCGGCAAGGACTTCGCGGCAACGCGGCGGCGGCTCTTTAC 1850
Db 376 GTTTCGTTTATGATACGCTTACGATTCGCGGATTAACGCGGCTTACGTTGTTTTC 435
Qy 1851 GCGCGCGCGCGGCGGAGGCGGCTGCGGAGGCTGCGGCTGACGTCGCGCGCGCTC 1910
Db 436 GCGCGCGCGGCGGAGGCGGCTG--CGTTGAGGTTGAGGTTGATTTGCTTCGTT 494
Qy 1911 GCGTACTACGCGGACCGCTCGGCGTGGGCGCGCGGAGTCCCGCAGTACTGCGGAC 1970
Db 495 GGTATTATACGTCGATTCGTCGCGGTTGGGCGTTGTTGATTTTGTGATTTGGGATTT 554
Qy 1971 AAGTCGCGCTGCTGCTGCTGCTGCTGCGGCAACGCGCGCGCGCGCGGATGGCC 2030
Db 555 AAGTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
Qy 2031 GCGCGCAATCCCTACCTGCGGAGAGGCGGAGGCGCTGCGCGGAGCGCTGCGCGCTG 2090
Db 614 GCGCTTATTTTATTTTATTTGCGGAGAGGTCAGAGGTTTGTGTCGAGCGTTCTGTTG 673
Qy 2091 CCGCCGCGCGCGCGCGAGAGCGCCAAAGCACTGTCGATTCAGCTGATCGAG 2150
Db 674 TCGTTC-CCGTCGTGAGAGCGTTAAGTTTGAATTTGTTCAATTTTGAATTTGATCGAG 732
Qy 2151 ACGCCCTCTCGATCAAGTCATCGACTCCAGCGACTCGGGGATTTAGAGAGCGCAAG 2210
Db 733 ACGTTTATTTTTCATTAAAGTTTATGATTTTATGCAATTCGCGGATTTAGATGTTAAG 792
Qy 2211 CGGAGCGGATCTGCGCGCGCGGAGACGCGCGTGTCCAGAGATTGTCGCCGCTCAAGAGC 2270
Db 793 CGGAGCGGATTTCTGCGTGTGATACGTTCTGTTTCAAGATTTGTTTCTTTAAAGAGC 852
Qy 2271 GAGGTGCTGCGCGAGCGGAGCTGCGAAGAAATGCGCAAGACATTAGCGGCTACTAT 2330
```

Db 853 GAGGTTGGTGTACCGGATGCGAGAGATTCGTTAAGGATATTAACGTTATAT 912
Qy 2331 GGCCTCTACTGCGACAGTAGAGCCGCCCTGCGCCGCCGCCGCCGAGACCC 2390
Db 913 GGTTTTATTCGTATAGTTAGTTCGTTTGTTCGTGCGTTTCGTCGCGGATTT 972
Qy 2391 CCAGCCAGCCCTCAACAGCTCTTCCCGAGCTCCGCTCCCACTCTCTTG 2444
Db 973 TTAGTTAGTTTATTAAGTTTATTTTATTTTATTTTATTTTATTTT 1026

RESULT 11
ID ACH15279 standard; cDNA; 457 BP.
XX ACH15279;
AC ACH15279;
DT 13-OCT-2003 (first entry)
XX
XX Human adult brain cDNA #2491.
DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KM genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
OS
XX US2003073623-A1.
XX 17-APR-2003.
PD 30-JUL-2001; 2001US-00918995.
PF 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
PR (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
PI
XX MPI; 2003-615964/58.
DR
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX
PS Claim 1; SEQ ID NO 2491; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docid=20030073623
XX
XX Sequence 457 BP; 81 A; 168 C; 85 G; 123 T; 0 U; 0 Other;

Query Match 13.2%; Score 419; DB 9; Length 457;

Best Local Similarity 98.8%; Pred. No. 6e-76;
Matches 422; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2299 AGAAGTGGCCCAAGACATTTAGCGGCTACTATGCTTCTACTGCGACAGTAGGCGGCC 2358
Db 31 AAACCTGGCCCAAGACATTTAGCGGCTACTATGCTTCTACTGCGACAGTAGGCGGCC 90
Qy 2359 CTGCG 2418
Db 91 CTACCG 150
Qy 2419 GCTCGCGCTCCCGACACACTCTCTCTGCGCGACCGACCTCATTTATTTGAACCTTGATGCGC 2478
Db 151 GCTCGCGCTCCCGACACACTCTCTCTGCGCGACCGACCTCATTTATTTGAACCTTGATGCGC 210
Qy 2479 GTCTGACGAGATATAGTGCAGGCTCTCGAGCGGTATTTAACTTTTGTGACAGCAGTC 2538
Db 211 GTCTGACGAGATATAGTGCAGGCTCTCGAGCGGTATTTAACTTTTGTGACAGCAGTC 270
Qy 2539 TCTGCAATTAGCTACCGACCTTCACTTGTGTAACCTTTGTTTCTTCTACTACT 2598
Db 271 TCTGCAATTAGCTACCGACCTTCACTTGTGTAACCTTTGTTTCTTCTACTACT 330
Qy 2599 CTCTCTCTGAGTATCTCTTACATTTCCCTCCCTCGTCTTCTCTTACTCTCT 2658
Db 331 CTCTCTCTGAGTATCTCTTACATTTCCCTCCCTCGTCTTCTCTTACTCTCT 390
Qy 2659 ACTTCTCTTCTTGTATGAACTCTTCACTTTAGAGACCTTGCGAGCTCTGTCAGGC 2718
Db 391 ACTTCTCTTCTTGTATGAACTCTTCACTTTAGAGACCTTGCGAGCTCTGTCAGGC 450
Qy 2719 AGCAGCG 2725
Db 451 AGCAGCG 457

RESULT 12
ID AAC00887 standard; cDNA; 388 BP.
XX AAC00887;
AC AAC00887;
DT 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 885.
DE Human
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX MPI; 2000-500381/45.
XX P-PSDB; AAG00881.
DR
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 885; 71pp + Sequence Listing; English.
PS
XX The present sequence is one of a large number of 5' ESTs derived from


```

QY      1392 ACTGAGTTCATGCGCGTACCGCTACCAAGAACAGGATATTACCAACTGAAATAGAT 1451
      878 ACGCAATTCATTCGAGTGAAGTCTCTACCAAAAACCGGATATTACCACTAAGATTTGAT 937
QY      1452 CACAAACCTTTTGGAAAAAGATTGGGATATTATGTGACAGATCTACACGGGCTGTAC 1511
      938 CATACACCTTTTGGAAAAAGCTTCAGAGCAACTATGATTCATATACACGCTTCAGAA 997
QY      1512 ATGACCGCGCTGACCCCTCGCCCAAGCACTCGCGCTCGCAGATCGTCCGCGGAGCC 1571
      998 AATGACAGGTTAACTTCATCTCCCAAGATTTCTTATGATCCCATCAGATTGTCCCTGGA 1057
Db      1572 CGCTACGCGCATGCGCGGCTCTTCTCTGACAGACAGTTCTGAGCACTACGCCAAGGCC 1631
      1058 GGTGCGTACGGGCTTCATCTCTTCCCGGAGCCCTTGTCAACACTTACTCAAGCC 1117
QY      1632 CGCTTCCACCGCGCGCGCGCGCGCGGCC 1660
      1118 CGCTATTATATATGCGAGAACCGTGCC 1146
Db

```

RESULT 14
ADP28768
ID ADP28768 standard, DNA; 3465 BP.
XX
AC ADP28768;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein encoding sequence SEQ ID #766.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; ds; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411032P.

```

PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411077P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 766; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosstatic,  
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein encoding sequence. The  
CC present sequence is available on WIPower and is not in the specification.  
XX  
XX Sequence 3465 BP; 847 A; 1005 C; 912 G; 701 T; 0 U; 0 Other;  
SQ  


```

Query Match 11.8%; Score 374.6; DB 12; Length 3465;
Best Local Similarity 68.8%; Pred. No. 1.3e-66;
Matches 515; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

```

QY      912 GCACAGGTATCTGTGCAACAGCGCCCTTGGCTGGAATTTACCGGACCAAGCGAG 971
      952 GCCACAGTTCATCTGCAACCGCGCTCTGUGCTCAATTCACCGCCCAACCTGAG 1011
QY      972 ATGATCATCAACCAAGGAGGAGGCAATGTTCTTTTAAGTTTAACTTTCTGGT 1031

```

```

Db      1012 ATGATCATTACGAAACAGGCGAGCGCATGTTCTTCTTCTTGAGCTTCAACATTAACGGA 1071
Qy      1032 CTGATCCCAAGCGCTCATATTAAATTTTGTGATGTGATTTGGCGGATCCCATAC 1091
Db      1072 CTCATATCCCACTGCGCATACATGTGTGTAGAGGTGTCTGGCGGACCCCAACAC 1131
Qy      1092 TGGAGGTTTCAAGGAGCAATGGGTCTCTTGCGGCAAGCGGACCAATGTGAAGA 1151
Db      1132 TGGCGCTTCCAGGGGCAATGGGTGTGACCTGTGCAAGCGGACCAATGACGAGGC 1191
Qy      1152 AATCGGCTTATATGATCCGATTCGCCCAACATGGGGCTCAGTGTAGCGGCAAGNA 1211
Db      1192 AACAAATGTATGTTCAACCAAGTCTCTTAATACGTGTCCACGTGATGACAGGAG 1251
Qy      1212 ATCTCTTTTGAATAATTAACCTTACGAACAACAAGAGCTTCAATATACATGGGAG 1271
Db      1252 ATTTCATTCGGGAAATTAATTAACCTACCAATACAAAGGCGCAATTAACAACACCGAG 1311
Qy      1272 ATGGTGGTTTTCAGTCTTGTGCAAGTACAGGCGCGCTGTCATGTGGTGAAGTGAAC 1331
Db      1312 ATGATAGTCTTACATCTTACACAAATACCAACCCGACGTGATATGTTGAAGTTACA 1371
Qy      1332 GAGGACGGGACGAGGAGCACTAGCCAGCCGCGCTGTCAGACGTTCACTTTCCTGAG 1391
Db      1372 GAGGATGGCGTGGAGGACTTGATGAGCCCTCAAGAACCCACATTTTACCTTCAGAA 1431
Qy      1392 ACTCATTTATGCGGTACACCGCTTACCAAGAACAGGATATTACAACTGAAATAGAT 1451
Db      1432 ACGCATTCATGTGAGTGACTGCTTACCAACCAACCGAATATTCTCACTAAAGATTGAT 1491
Qy      1452 CACAACCCCTTTGCAAAAGGATTTGGGGATTAATTATGACAGATCTACACCGGCTGTGAC 1511
Db      1492 CATTAACCCCTTTGCAAAAGGCTTACAGACAACTATGATTCATGTACACCGCTTCAGAA 1551
Qy      1512 ATGACCGGCTGACCCCTGCGCCAAAGACTGCGCGCTGTCAGATGCTGCGCGGAGCC 1571
Db      1552 AATGACAGGTAACTGCATCTCCAGAGATTCTCGTAGATCCCATCAATGATGTCCTGGA 1611
Qy      1572 CGCTACGCGCATGCGCGCTTTTCTGTGACGACAGCACTGTGAGCACTACGCGCAAGCC 1631
Db      1612 GGTGCGTACGCGGTCAATCTTCTTCCGAGGCCCTTGTCAACACTTTACCTCAAGGC 1671
Qy      1632 GCGTTCCACCGCGGCGCGGCGCGGAGCC 1660
Db      1672 CGCTATTATATGCGAGAGAACCGTGCC 1700

```

RESULT 15

AA039695
ID AA039695 strand; DNA; 389 BP.

```

AC      AA039695;
XX      25-MAR-2003 (revised)
DT      20-MAY-1993 (first entry)
XX      Expressed Sequence Tag human gene marker EST00360.
DE      Expressed Sequence Tag human gene marker EST00360.
KW      expressed sequence tag; human genome project; chromosome;
KW      human gene sequencing; PCR mapping; somatic cell hybrids;
KW      sublocalisation; gene tagging; tissue typing.
XX      Synthetic.
OS      Synthetic.
XX      WO9300353-A1.
XX      07-JAN-1993.
XX      19-JUN-1992; 92WO-US005222.
XX      20-JUN-1991; 91US-00716831.
XX      12-FEB-1992; 92US-00837195.
PR

```

```

XX      (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA      Venter JC, Adams MD;
XX      WPI; 1993-036325/04.
XX      Particular expressed sequence tags from human CDNA - corresponds to
PT      transcription prods. of genes, useful for tagging genes, mapping
PT      chromosomes and tissue typing.
XX      Claim 3; Page 105; 199pp; English.
XX      This sequence represents an EST (expressed sequence tag) ESTs are markers
CC      for human genes actually transcribed in vivo. Unlike the random genomic
CC      DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.
CC      The use of ESTs could facilitate the tagging of most expressed human
CC      genes within a few years at a fraction of the cost of complete genomic
CC      sequencing. Using PCR primers AA039419-039580 (sequences designed from
CC      the ESTs) sublocalisation of an EST can be achieved with panels of
CC      fragments from specific chromosomes or pools of large genomic clones in
CC      an analogous manner. This sequence represents EST00360. (Updated on 25-
CC      Mar-2003 to correct PN field.)
XX      SQ      Sequence 389 BP; 97 A; 81 C; 71 G; 139 T; 0 U; 1 Other;
XX
Query Match      11.4%; Score 363; DB 2; Length 389;
Best Local Similarity 99.2%; Pred. No. 1.7e-64;
Matches 385; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy      2775 TCTAGAACCTTGACCCACCCAGCGCGCTTCTTATCCCGAGTGAATGAGATGGA 2834
Db      1 TCTAGAACCTTGACCCACCCAGCGCGCTTCTTATCCCGAGTGAATGAGATGGA 60
Qy      2835 TGAATGAGGAGTGTATATATTTTATGTGGAACAAGCCTGTGAATGATTGTACATAG 2894
Db      61 TGAATGAGGAGTGTATATATTTTATGTGGAACAAGCCTGTGAATGATTGTACATAG 120
Qy      2895 TGTATATTTATTTGAGGAATGGCTAGTTTATCTGTGCAAGGCACAAAACGAGTTCA 2954
Db      121 TGTATATTTATTTGAGGAATGGCTAGTTTATCTGTGCAAGGCACAAAACGAGTTCA 180
Qy      2955 TGCTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 3014
Db      181 TGCTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 240
Qy      3015 TTCTCTCTCTTTTATTTTCTTGTGAGATATATTTCTTAAGGCTCTAGAAACATGAAT 3074
Db      241 TTCTCTCTCTTTTATTTTCTTGTGAGATATATTTCTTAAGGCTCTAGAAACATGAAT 300
Qy      3075 ACTGAGTAGT-GATGGGTTTCCACTCTCTCAATCGGTGAGTAATAATTACTAT- 3132
Db      301 ACTGAGTAGTGAATGGGTTTCCACTCTCTCAATCGGTGAGTAATAATTACTATG 360
Qy      3133 GTGCCCTAATGACACAATAATAGCTAAG 3160
Db      361 GTGCCCTAATGACACAATAATAGCTAAG 388

```

Search completed: December 20, 2005, 08:14:21
Job time : 1811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 07:07:03 ; Search time 12029 Seconds
(without alignments)
12392.023 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....cacccaacaccttaag 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hic: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gsa1: *
10: gb_gsa2: *
11: gb_gsa3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1002.6	31.5	1116	5	BX418703	BX418703 BX418703
2	935.4	29.4	957	5	BX441549	BX441549 BX441549
3	897	28.2	921	5	BX420367	BX420367 BX420367
4	823	25.8	930	5	BX1912693	BX1912693 603176344
5	811.4	25.5	824	5	BX418656	BX418656 BX418656
6	809	25.4	820	5	BX1915063	BX1915063 603177258
7	801.6	25.2	872	3	BX1914330	BX1914330 603182311
8	798.2	25.1	1169	3	BQ067178	BQ067178 AGENCOURT
9	789.8	24.8	1044	3	BM922065	BM922065 AGENCOURT
10	766.6	24.1	945	3	BX1914132	BX1914132 603182077
11	761	23.9	966	3	BX1754056	BX1754056 603027611
12	760	23.9	831	1	AM003023	AM003023 wq61d04.x
13	749.6	23.5	910	7	CN643299	CN643299 ILLUMIGEN
14	722.4	22.7	768	8	DR001136	DR001136 TC115971
15	720	22.6	910	7	CN805179	CN805179 ILLUMIGEN
16	716	22.5	748	3	BX1917426	BX1917426 603184973
17	710	22.3	991	3	BX1757271	BX1757271 603030777
18	699.8	22.0	791	3	BX1918360	BX1918360 603182871
19	695.6	21.8	779	3	BX1820633	BX1820633 603034884
20	693.4	21.8	740	3	BX1917781	BX1917781 603183881
21	692.8	21.7	777	6	CD355525	CD355525 UT-M-FY0-
22	691.4	21.7	1154	3	BM807553	BM807553 AGENCOURT

23	687.6	21.6	772	3	BI823704	BI823704 603040821
24	682.4	21.4	739	3	BI918597	BI918597 603176527
25	681.8	21.4	1028	3	BM548197	BM548197 AGENCOURT
26	664.8	20.9	712	6	CA326858	CA326858 UI-M-FY0-
27	664.6	20.9	749	5	BQ443529	BQ443529 UI-M-FY0-
28	664.6	20.7	787	3	BI917926	BI917926 603181686
29	658.4	20.7	787	3	CA326808	CA326808 UI-M-FY0-
30	642.2	20.2	716	6	CA315364	CA315364 UI-M-FY0-
31	640.8	20.1	716	6	BO179248	BO179248 UI-M-FY0-
32	637.6	20.0	716	3	BO179248	BO179248 UI-M-FY0-
33	625	19.6	674	6	CB244022	CB244022 UI-M-FY0-
34	616	19.3	639	3	BI819019	BI819019 603033141
35	613	19.2	765	3	BI912695	BI912695 603176346
36	610.6	19.2	1015	3	BM921945	BM921945 AGENCOURT
37	592.2	18.6	606	1	AM590019	AM590019 hg29e09.x
38	585.4	18.4	700	6	CF531585	CF531585 UI-M-FY0-
39	582.2	18.3	801	8	CB527374	CB527374 UI-M-FY0-
40	581	18.2	583	3	BP228857	BP228857 BP228857
41	560.4	17.6	582	3	BP228857	BP228857 BP228857
42	554.8	17.4	855	8	CB829223	CB829223 JGI CAAR6
43	515.8	16.2	534	5	BX114402	BX114402 BX114402
44	508.8	16.0	541	1	AI201654	AI201654 q141f08.x
45	506.8	15.9	566	7	CV562445	CV562445 UI-M-FY0-

ALIGNMENTS

RESULT 1
BX418703
LOCUS
DEFINITION
CS0DF008YF20 5-PRIME, mRNA sequence.
ACCESSION
BX418703.2 GI:46957805
VERSION
BX418703
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1116)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 15, 2003 this sequence version replaced gi:30769503.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DF008YF20&cl=1851.f.
Location/Qualifiers
1..1116
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF008YF20"
/issue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1ib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

FEATURES

ORIGIN

188 GGGGCTCTGAGCTTCTGTCAGATTTCATGATGCTCTGTCGAGATCGTACTCTCTCT 247
 QY 627 CAGTCCAGCCAGCAGCAGCTCTGCGGCACTGCTCCAGATGTCATGTTCCGTTACCCGCG 686
 Db 248 CAGTCCAGCCAGCAGCAGCTCTGCGGCACTGCTCCAGATGTCATGTTCCGTTACCCGCG 307
 QY 687 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 746
 Db 308 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 367
 QY 747 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 806
 Db 368 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 426
 QY 807 TACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 866
 Db 427 TACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 486
 QY 867 TTTCTACAGCTTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 926
 Db 487 TTTCTACAGCTTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 546
 QY 927 TGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 986
 Db 547 TGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 606
 QY 987 CAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1046
 Db 607 CAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 666
 QY 1047 CATTAACAATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
 Db 667 CATTAACAATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
 QY 1107 GCGAATAGGCTCTCTGCGGCAAGCGGCAACCAATGCGAAGGAAATGCGGCTATATG 1166
 Db 727 GCGAATAGGCTCTCTGCGGCAAGCGGCAACCAATGCGAAGGAAATGCGGCTATATG 786
 QY 1167 CATCCGAGTTCCTCCCAACAAGCGGCTCTGCGGCAAGCGGCAAGCGGCTATATG 1226
 Db 787 CATCCGAGTTCCTCCCAACAAGCGGCTCTGCGGCAAGCGGCTATATG 846
 QY 1227 TTTAAACTTACGAAACAACAAGCGGCTCTGCGGCAAGCGGCTATATG 1286
 Db 847 TTTAAACTTACGAAACAACAAGCGGCTCTGCGGCAAGCGGCTATATG 906
 QY 1287 TCTTGGCAAGATGACAGCGGCTCTGCGGCAAGCGGCTATATG 1337
 Db 907 TCTTGGCAAGATGACAGCGGCTCTGCGGCAAGCGGCTATATG 957

RESULT 3
 BX420367 921 bp mRNA linear EST 01-MAY-2004
 LOCUS BX420367 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 DEFINITION CS0DF027YH06 5-PRIME, mRNA sequence.
 ACCESSION BX420367
 VERSION BX420367
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 921)
 11, W.B., Gruber, C., Jesse, J., and Polyes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30650954.
 Contact: Genoscope
 Genoscope - Centre National de Sequenage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 1851.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?cs=CS0DF027DD03QP1&c=1851.f.
 Location/Qualifiers
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF027YH06"
 /release_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_id="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN
 Query Match 28.2%; Score 897; DB 5; Length 921;
 Best Local Similarity 99.7%; Pred. No. 7.2e-206;
 Matches 919; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 QY 358 ATGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 417
 Db 1 ATGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
 QY 418 TTAATCGACCACTGACCACTGAGAGAAATTTCACTTTGAAAATAATTCACGAGGAG 477
 Db 61 TTAATCGACCACTGACCACTGAGAGAAATTTCACTTTGAAAATAATTCACGAGGAG 120
 QY 478 TGAAGCAATCAGTCAATGACCAATTTCTCTGACTGACCTCAAGAGACTCACAGGAGAG 537
 Db 121 TGAAGCAATCAGTCAATGACCAATTTCTCTGACTGACCTCAAGAGACTCACAGGAGAG 180
 QY 538 AGAGAAATGAACT 597
 Db 181 AGAGAAATGAACT 240
 QY 598 GCTCTGTCAGATGCTACT 657
 Db 241 GCTCTGTCAGATGCTACT 300
 QY 658 CTCCAGATGCAATGTTCCCGTACCCCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 717
 Db 301 CTCCAGATGCAATGTTCCCGTACCCCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
 QY 718 TCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 777
 Db 361 TCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
 QY 778 GCTCTCTGTCAGATGCTACT 837
 Db 421 GCTCTCTGTCAGATGCTACT 479
 QY 838 AGTACGCGCAGCTCTACCAAGAGAGCTCGGTTTACAGTTTCTCTCAACCAAGCGGAG 897
 Db 480 AGTACGCGCAGCTCTACCAAGAGAGCTCGGTTTACAGTTTCTCTCAACCAAGCGGAG 539
 QY 898 TGTGTCGCGGCAAG 957
 Db 540 TGTGTCGCGGCAAG 599
 QY 958 GGCACCAACGAGAGATGATCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
 Db 600 GGCACCAACGAGAGATGATCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
 QY 1018 TTAACAATTTCTGTCTGATCCAGCGGCTCATTAACAATATTTTGTGATGATGATTTGG 1077

Db 660 TTAACATTCTGTCATCCACGGCTCATTAATATTTGTGATGATTTGG 719
Qy 1078 CGGATCCCAATCTGAGGTTTCAAGAGCAATGGGTTCTTGGCGCAAGCGACA 1137
Db 720 CGGATCCCAATCTGAGGTTTCAAGAGCAATGGGTTCTTGGCGCAAGCGACA 779
Qy 1138 CCATGTCGAAGAAATGGGCTTATGCAATCCGATTCGCCCAACCTGGGGTCACT 1197
Db 780 CCAATGTCGAAGAAATGGGCTTATGCAATCCGATTCGCCCAACCTGGGGTCACT 839
Qy 1198 GGATGCGCCAGAAATCTC-TTTTGGAAATTTAAATTCAGAACAAAGAGCTTCA 1256
Db 840 GGATGCGCCAGAAATCTCTTTTGGAAATTTAAATTCAGAACAAAGAGCTTCA 899
Qy 1257 AATAACAATGGCGAGATGTTG 1278
Db 900 AATAACAATGGCGAGATGTTG 921

RESULT 4
B1912693 930 bp mRNA linear EST 16-OCT-2001
LOCUS 603176344F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240506 5',
DEFINITION mRNA sequence.
B1912693
ACCESSION B1912693.1 GI:16176907
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 930)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: LAM11606 row: h column: 11
High quality sequence stop: 810.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240506"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 25.8%; Score 823; DB 3; Length 930;
Best Local Similarity 95.7%; Pred. No. 6, 2e-189;
Matches 889; Conservative 0; Mismatches 35; Indels 5; Gaps 4;
279 GCGAGTGTTCAGGTTCTAGAGCTATGACGCTGACACTGCTTCTCTCTATCATG 338

Db 1 GCGAGTGTTCAGGTTCTAGAGCTATGACACTGAGAGACTGCTTCTCTCTATCATG 60
Qy 339 CTCCTCCAGAAATTTCTCAATGTGAGAGAGAGTACCAATTCAGGCGGATCCGAGCTT 398
Db 61 CTCCTCCAGAAATTTCTCAATGTGAGAGAGAGTACCAATTCAGGCGGATCCGAGCTT 120
Qy 399 GTCTTGACAGATCAATCCCATTTATCTGACCACTGACCAACCTGAGAGAAATTCCTTTG 458
Db 121 GTCTTGACAGATCAATCCCATTTATCTGACCACTGACCAACCTGAGAGAAATTCCTTTG 180
Qy 459 AAAAAATTAACAGAGGAGATGACGAATCAGTCAATACAGCAATTTCTGACTCCAAG 518
Db 181 AAAAAATTAACA-GGGAGTAGCAATCAATCAGATCAGATCAATTTCTGACTCCAAG 239
Qy 519 GACTCACCAGAGGAGCGTCCAGAGAAATCTCTCTGCTTGGAGGGGCTCTGAG 578
Db 240 GACTCACCAGAGGAGCGTCCAGAGAAATCTCTCTGCTTGGAGGGGCTCTGAG 299
Qy 579 CTTGCTCAGATTCGATGCTCTGCTGACAGATCGTCACTCTCTCTGAGTCCAGCCAG 638
Db 300 CTTGCTCAGATTCGATGCTCTGCTGACAGATCGTCACTCTCTCTGAGTCCAGCCAG 359
Qy 639 CCACAGTCTGCGGCTGCTCTCCAGTGCATGTTCCCGTACCCGCGCAGACGACCG 698
Db 360 CCACAGTCTGCGGCTGCTCTCCAGTGCATGTTCCCGTACCCGCGCAGACGACCG 419
Qy 699 GCGACACCCGCTTCTCATTCGAGCGCTTACCGGCTCAATGAGCCACACCCGCGCTATC 758
Db 420 GCGACACCCGCTTCTCATTCGAGCGCTTACCGGCTCAATGAGCCACACCCGCGCTATC 478
Qy 759 ACCAAGGAGGCTTCAACAGAGCTCTGTCCAATCTCTGCGCGAGGATACCCACGCGCC 818
Db 479 ACCAAGGAGGCTTCAACAGAGCTCTGTCCAATCTCTGCGCGAGGATACCCACGCGCC 538
Qy 819 GGTACCTCTTACCAACAGAGTACGAGCTCTCTTACCAAGAGCTCTCTTACCAATTC 878
Db 539 GGTACCTCTTACCAACAGAGTACGAGCTCTCTTACCAAGAGCTCTCTTACCAATTC 598
Qy 879 TCCCTCCACCCAGCGCGGGCTGTGTCCCGGCAACACAGGATGCTGTGACACAGGCC 938
Db 599 TCCCTCCACCCAGCGCGGGCTGTGTCCCGGCAACACAGGATGCTGTGACACAGGCC 658
Qy 939 CTTTGCGTGAATTTTCCACCGGACCAACGAGATGATCATCAACCAACGAGGAGCGC 998
Db 659 CTTTGCGTGAATTTTCCACCGGACCAACGAGATGATCATCAACCAACGAGGAGCGC 718
Qy 999 ATGTTTCTTTTAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1058
Db 719 ATGTTTCTTTTAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 778
Qy 1059 TTTTGATGATGATTTTGGCGGATCCCAATCACTGAGGATTTTCAAGGAGGCAATGGGTT 1118
Db 779 TGTGGAATGTGAATTTGGCGGATCCCAATCACTGAGGATTTTCAAGGAGGCAATGGGTT 837
Qy 1119 CTTTGCGGCAAGGAGGACCAATGTGCAAGAAATCGGCTTATGATCCGATTTCC 1178
Db 838 TCTTGCGGCAAGGAGGACCAATGTGCAAGAAATCGGCTTATGATCCGATTTCC 897
Qy 1179 CCACACTGGGGCTCACTGATCGCCA 1207
Db 898 CC-AAATGGGGCTCACTGATCGCCA 924

RESULT 5
BX418656 824 bp mRNA linear EST 03-MAY-2004
LOCUS BX418656 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CSDDF008Y10 5-PRIME, mRNA sequence.
ACCESSION BX418656
VERSION BX418656.2 GI:46954137
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 824)
TITLE	L1, W.B., Gruber, C., Jeesee, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001) On May 15, 2003 this sequence version replaced gi:30765811. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1851.f For more information about this cluster, see http://www.genoscope.cns.fr/cdna?c=CS0DP008BD05QPl&c=1851.f.
FEATURES	Location/Qualifiers
source	1..824 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DP008BG10" /tissue_type="FETAL BRAIN" /dev_stage="fetal" /clone_1ib="Homo sapiens FETAL BRAIN" /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN	
Query Match	25.5%; Score 811.4; DB 5; Length 824;
Best Local Similarity	99.8%; Pred. No. 3.9e-185;
Matches	823; Conservative 0; Mismatches 1; Indels 1; Gaps 1
Db	200 GAAGTGCCTTCTGTCTAGTAGAGGGGCTGTGAGATTCTTCTAGTTATGATTAATAGACTT 259 1 GATGTGCTTCTGTCTAGTAGAGGGGCTGTGAGATTCTTCTAGTTATGATTAATAGACTT 60
Qy	260 TAAAAACACGGGACGGGAGGGGCGAGTTCACAGTTCTAAGCTATGACAGCTGAGACTG 319
Db	61 TAAAAACACGGGACGGGAGGGGCGAGTTCACAGTTCTAAGCTATGACAGCTGAGACTG 120
Qy	320 CCTTCTCCTTCTATCATGCTCTCCAAAGAAATTTCTCAATGTGAGCAGCAGCTACCA 379 121 CCTTCTCCTTCTATCATGCTCTCCAAAGAAATTTCTCAATGTGAGCAGCAGCTACCA 180
Db	380 TTCAAGCGGATCGAGACTTGTCTTGACAGATCATCCATTATCTGACCACTGACAACCT 439 181 TTCAAGCGGATCGAGACTTGTCTTGACAGATCATCCATTATCTGACCACTGACAACCT 240
Qy	440 GGAGAGAAGTTCACCTTTGAAAAAAATTATCCAGGGGGATGAGAAATCACTCAGATTACA 499
Db	241 GGAGAGAAGTTCACCTTTGAAAAAAATTATCCAGGGGGATGAGAAATCACTCAGATTACA 300
Qy	500 CAATTTTCTGACTCCAGAGACTCACACAGGGGACGTCACAGAGAAATTAATCTCTCTGCT 559 301 CAATTTTCTGACTCCAGAGACTCACACAGGGGACGTCACAGAGAAATTAATCTCTCTGCT 360
Db	560 CTTGACGGGGTCTGAGCTTGTGCAAGTTTGATGCGCTGCTGCAAGATCGTAACCT 619 361 CTTGACGGGGTCTGAGCTTGTGCAAGTTTGATGCGCTGCTGCAAGATCGTAACCT 420
Qy	620 CCTCTCTCACTCAGACGACGACAGCTGCGGGCACTGCTCCAGAGGCAATGTTCCCGTA 679 421 CCTCTCTCACTCAGACGACGACAGCTGCGGGCACTGCTCCAGAGGCAATGTTCCCGTA 480

QY	680	CCCCGGGCGAGGAGGAGCCGCGGCTTCTCTCATTCGCGAGCCGTACAT	739
QY	680	CCCCGGGCGAGGAGGAGCCGCGGCTTCTCTCATTCGCGAGCCGTACAT <td>739</td>	739
Db	481	CCCCGGGCGAGGAGGAGCCGCGGCTTCTCTCATTCGCGAGCCGTACAT	540
QY	740	GGCCCAACCAACCGGTCATCAACCAAGGAGCTTACACAGCTTCTCTCAATCTTCGCC	759
Db	541	GGCCCAACCAACCGGTCATCAACCAAGGAGCTTACACAGCTTCTCTCAATCTTCGCC	600
QY	800	GCAAGGAGTATCCCAACGCGCGGCTTACCCCTTACCCACAGAGTACGGGCACTTCACCAAGG	859
Db	601	GCAAGGAGTATCCCAACGCGCGGCTTACCCCTTACCCACAGAGTACGGGCACTTCACCAAGG	659
QY	860	AGCTCGGTTCTTACCAAGTTCTCTCTTCCACCAAGCCGCGGCTTCGCCGCAAGCAGAGT	919
Db	660	AGCTCGGTTCTTACCAAGTTCTCTCTTCCACCAAGCCGCGGCTTCGCCGCAAGCAGAGT	719
QY	920	GTACCTGTGACAGAGCCCTTGTGCTGAATTTTACCCGACCAACGAGATGATCAT	979
Db	720	GTACCTGTGACAGAGCCCTTGTGCTGAATTTTACCCGACCAACGAGATGATCAT	779
QY	980	CACCAACAGGGAAGGCGATGTTCTCTTTTAAAGTTTAAAGT	1024
Db	780	CACCAACAGGGAAGGCGATGTTCTCTTTTAAAGTTTAAAGT	824

RESULT 6

BI915063

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

820 bp

mRNA

linear

EST 16-OCT-2001

603177258F1 NIH_MGC_121

Homo sapiens

CDNA clone IMAGE:5241542

5'

BI915063

mRNA sequence.

BI915063.1

GI:16199126

EST.

Homo sapiens (human)

Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 820)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNase Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Plate: L1AM1609

row: C

column: 15

High quality sequence

stop: 816.

Location/Qualifiers

1..820

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5241542"

/lab_host="DH10B"

/clone_id="NIH_MGC_121"

/note="Organ: brain; Vector: pCMV-Sport5; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC library."

ORIGIN

Query Match 25.4%; Score 809; DB 3; Length 820;
 Best Local Similarity 99.9%; Pred. No. 1.5e-184;
 Matches 820; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 221 GGGGGCTGTGATTTCTAGATTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGG 280
 DB 1 GGGGGCTGTGATTTCTAGATTATGATTAATAGACTTTAAAAACAGGAGCGGAGCGG 59

QY 281 CGAGTGTTCAGGTTCTAGAGCTATGACGTGAGAGCACTGCTTTCTCTCTATATGCT 340
 DB 60 CGAGTGTTCAGGTTCTAGAGCTATGACGTGAGAGCACTGCTTTCTCTCTATATGCT 119

QY 341 CTCGAGAAATTTCTCAATGTGAGAGCACTGAGCACTTCCAGGCGGATCCGACTTGT 400
 DB 120 CTCGAGAAATTTCTCAATGTGAGAGCACTGAGCACTTCCAGGCGGATCCGACTTGT 179

QY 401 CTCGAGCAATCCCATTTATCTCGACCACTGACAACTGAGAGAAAGTTCACTTTGAA 460
 DB 180 CTCGAGCAATCCCATTTATCTCGACCACTGACAACTGAGAGAAAGTTCACTTTGAA 239

QY 461 AAAAATTCAGAGGGGATGACGAATCATGACATATACGACATTTTCTCTGATCCAGAGA 520
 DB 240 AAAAATTCAGAGGGGATGACGAATCATGACATATACGACATTTTCTCTGATCCAGAGA 299

QY 521 CTCACGAGGGGACGTCACAGAGATTAATCTCTCTCTCTGAGACGGGGTCTTGAGCT 580
 DB 300 CTCACGAGGGGACGTCACAGAGATTAATCTCTCTCTCTGAGACGGGGTCTTGAGCT 359

QY 581 TCGTCAAGATTTGATGAGCTCTGCTGCTGACAGATGCTACTCTCTCTCTGAGTCCAGCC 640
 DB 360 TCGTCAAGATTTGATGAGCTCTGCTGCTGACAGATGCTACTCTCTCTCTGAGTCCAGCC 419

QY 641 AAGCTGTGGGCGACGCTCCGAGGCAATGTTCCGTAACCCCGGACAGAGACGGCC 700
 DB 420 AAGCTGTGGGCGACGCTCCGAGGCAATGTTCCGTAACCCCGGACAGAGACGGCC 479

QY 701 GCACCCCGCTTCTCCATGCGAGCGCTTACGCGCTACATGAGCCCAACCCCGTATCATC 760
 DB 480 GCACCCCGCTTCTCCATGCGAGCGCTTACGCGCTTACGCGCCCAACCCCGTATCATC 539

QY 761 CAACGAGGCTTCAACAGGCTCTCTGTCACACTCTTCCGCGAGGAGTACCCCAAGCGCG 820
 DB 540 CAACGAGGCTTCAACAGGCTCTCTGTCACACTCTTCCGCGAGGAGTACCCCAAGCGCG 599

QY 821 CTAACCCCTTACCAAGCACTTACGCGCTTACGCGCTTACGCGCTTACGCGCTTACGCG 880
 DB 600 CTAACCCCTTACCAAGCACTTACGCGCTTACGCGCTTACGCGCTTACGCGCTTACGCG 659

QY 881 CTCACCCAGCGGGGCTGTGCTCCGCGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCT 940
 DB 660 CTCACCCAGCGGGGCTGTGCTCCGCGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCT 719

QY 941 TTGGCTGAATTTTACCGGACCAACGAGATGATCATACCAACAGGGAAGGCGCAT 1000
 DB 720 TTGGCTGAATTTTACCGGACCAACGAGATGATCATACCAACAGGGAAGGCGCAT 779

QY 1001 GTTTCCTTTTAAAGTTTAAATTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
 DB 780 GTTTCCTTTTAAAGTTTAAATTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 820

RESULT 7
 BI914330 872 bp mRNA 1linear EST 17-OCT-2001
 LOCUS BI914330
 DEFINITION 60318231fl NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246210 5',
 mRNA sequence.
 ACCESSION BI914330
 VERSION BI914330.1 GI:16198838
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 872)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.lnl.gov
 Plate: LMNL1621 row: F column: 03
 High quality sequence stop: 853.

FEATURES
 source
 Location/Qualifiers
 1..872
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246210"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 25.2%; Score 801.6; DB 3; Length 872;
 Best Local Similarity 97.7%; Pred. No. 9.4e-183;
 Matches 856; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

QY 367 GCAGTACCCATTTACAGGCGGATCCGAGCTTGTCTTGACAGATCATCTTATCTCGA 426
 DB 1 GCAGTACCCATTTACAGGCGGATCCGAGCTTGTCTTGACAGATCATCTTATCTCGA 60

QY 427 CCACGACACCTGAGAGAGTTCACCTTTGAAAAAATTAACAGGGGAGTGAATC 486
 DB 61 CCACGACACCTGAGAGAGTTCACCTTTGAAAAA--TTACAGGGAGTGAATC 118

QY 487 AGTCAGATACAGACATTTTCTGACTCCAGAGCTACCAAGGGAGCGTCCAGAGAAATA 546
 DB 119 AGTCAGATACAGACATTTTCTGACTCCAGAGCTACCAAGGGAGCGTCCAGAGAAATA 178

QY 547 AACTCTCTCTGTGTGGAGCGGGTCTGTAGCTTCTGACAGTTTGGATGGCTCTGTG 606
 DB 179 AACTCTCTCTGTGTGGAGCGGGTCTGTAGCTTCTGACAGTTTGGATGGCTCTGTG 238

QY 607 CAGATCGTACTCTCTCTGAGTCAGGCGGACAGTCTGCGGCACTGCTCCAGT 666
 DB 239 CAGATCGTACTCTCTCTGAGTCAGGCGGACAGTCTGCGGCACTGCTCCAGT 298

QY 667 CCATGTTCCGTTACCCCGGCGAGCACCGGCGCACCCCGCTTCTCATCGGACGCC 726
 DB 299 CCATGTTCCGTTACCCCGGCGAGCACCGGCGCACCCCGCTTCTCATCGGACGCC 358

QY 727 CTAGCGCTTACATGGCCCAACCGGCTCATCCCAAGAGCTTACCAAGCTTCTGT 786
 DB 359 CTAGCGCTTACATGGCCCAACCGGCTCATCCCAAGAGCTTACCAAGCTTCTGT 418

QY 787 CCAACTCTCGCGGAGGATACCCCAAGCGCGGCTACCCCTTACCCACAGAGTACGCC 846
 DB 419 CCAACTCTCGCGGAGGATACCCCAAGCGCGGCTACCCCTTACCCACAGAGTACGCC 478

QY 847 ACTCTACCAAGAGCTCCGTTTACAGATTCTCTTCCACCCAGCCGGGCGTGTGCCG 906

Db 479 ACTCTTACCAAGAGTCCGTTCTACCAATCTCTCCACCAACCGGCGTGTGCGG 538
 Qy 907 GCANAAGCAGAGTGTACCTGTCCAAACAGCCCTTTGGCTGAATTTCAACGGCAGCAAA 966
 Db 539 GCANAAGCAGAGTGTACCTGTCCAAACAGCCCTTTGGCTGAATTTCAACGGCAGCAAA 598
 Qy 967 CGGAGATGATCATCACCAACAGGAGCGCATGTTCTTTTAAAGTTTAAATTT 1026
 Db 599 CGGAGATGATCATCACCAACAGGAGCGCATGTTCTTTTAAAGTTTAAATTT 658
 Qy 1027 CTGGCTGTGATCCGCTGCTTCAATATTTTGTGATGATGATTTGGCGATCCCA 1086
 Db 659 CTGGCTGTGATCCGCTGCTTCAATATTTTGTGATGATGATTTGGCGATCCCA 718
 Qy 1087 ATCACTGAGAGTTTCAAGAGCAATGGGTTCTTGGCGCAAGCGGACCAATGTGC 1146
 Db 719 ATCACTGAGAGTTTCAAGAGCAATGGGTTCTTGGCGCAAGCGGACCAATGTGC 778
 Qy 1147 AAGGAAA--TCGGGTCTATATGATCGGATTCCTTGGCGCAAGCGGACCAATGTGC 1204
 Db 779 AAGGAAAATCGGGGTCTATATGATCGGATTCCTTGGCGCAAGCGGACCAATGTGC 837
 Qy 1205 CCAAGAAATCTCTTTGGAATAATTAAGTTAGCA 1240
 Db 838 -CAAGAAATCTCTTTGGAATAATTAAGTTAGCA 972

RESULT 8
 LOCUS B0067178 1169 bp mRNA linear EST 02-APR-2002
 DEFINITION AGENCOURT_6767481 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5751547
 5', mRNA sequence.
 ACCESSION B0067178
 VERSION B0067178.1 GI:19896224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 1169)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12784 row: e column: 20
 High quality sequence start: 16
 High quality sequence stop: 685.
 Location/Qualifiers
 1..1169
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5751547"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain; lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. library is normalized and
 enriched for full-length clones and was constructed by C.

FEATURES
 source

Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 Query Match 25.1%; Score 798.2; DB 3; Length 1169;
 Best Local Similarity 96.4%; Pred. No. 6.7e-182;
 Matches 826; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

Qy 826 CCTACCCACAGCATGACCGCCCACTTACCAAGAGAGTCCGTTCTACCAAGTTTCCTCA 885
 Db 36 CCTACCCACAGCATGACCGCCCACTTACCAAGAGAGTCCGTTCTACCAAGTTTCCTCA 95
 Qy 886 CCCAGCCGGGCTGTGCTCCGCAAGAGCAGTGTACTGTGCAACAGGCCCTTTGGC 945
 Db 96 CCCAGCCGGGCTGTGCTCCGCAAGAGCAGTGTACTGTGCAACAGGCCCTTTGGC 155
 Qy 946 TGAATTTTACCGGACCAACCGAGATGATCATCACCAACAGGAGAGCGCATGTTTC 1005
 Db 156 TGAATTTTACCGGACCAACCGAGATGATCATCACCAACAGGAGAGCGCATGTTTC 215
 Qy 1006 CTTTTTAAAGTTTAACTTCTGTCTGATTCCTCAGGCTCATTAACAATTTTGTGG 1065
 Db 216 CTTTTTAAAGTTTAACTTCTGTCTGATTCCTCAGGCTCATTAACAATTTTGTGG 275
 Qy 1066 ATGTGATTTTGGCGGATCCCAATCATCTGAGAGTTTCAAGAGCAATGGTTCCTTGC 1125
 Db 276 ATGTGATTTTGGCGGATCCCAATCATCTGAGAGTTTCAAGAGCAATGGTTCCTTGC 335
 Qy 1126 GCANAAGCAGACCAATGTGCAAGGAAATCGGGTCTATATGATCGGATTCCTTGC 1185
 Db 336 GCANAAGCAGACCAATGTGCAAGGAAATCGGGTCTATATGATCGGATTCCTTGC 395
 Qy 1186 CTGGGGCTCATCTGATGCGCCCAAGAACTCTTTTGGAAATTAACCTTGCAGCAACA 1245
 Db 396 CTGGGGCTCATCTGATGCGCCCAAGAACTCTTTTGGAAATTAACCTTGCAGCAACA 455
 Qy 1246 AAGGAGCTTCAAAATTAACATGGGCGAGATGATGTTTCAAGTCTTGCAGATACAGC 1305
 Db 456 AAGGAGCTTCAAAATTAACATGGGCGAGATGATGTTTCAAGTCTTGCAGATACAGC 515
 Qy 1306 CCCGCTCATGCTGTGTGAAAGTGAACGAGACGCGACGAGACACTAGCCGCGGC 1365
 Db 516 CCCGCTCATGCTGTGTGAAAGTGAACGAGACGCGACGAGACACTAGCCGCGGC 575
 Qy 1366 GCGTCAAGCGTTCACTTCTTCTGAGACTCATGCGGTCAACCGCTTACCAACA 1425
 Db 576 GCGTCAAGCGTTCACTTCTTCTGAGACTCATGCGGTCAACCGCTTACCAACA 635
 Qy 1426 CGGATTTAACAATGAATGATCAACCCCTTTGCAAAAGATTTGGGATTAAT 1485
 Db 636 CGGATTTAACAATGAATGATCAACCCCTTTGCAAAAGATTTGGGATTAAT 695
 Qy 1486 ATGACAGCATCTACACCGGCTGTGATGACCGGCTGACCCCTTGCAGCACTGCG 1545
 Db 696 ATGACAGCATCTACACCGGCTGTGATGACCGGCTGACCCCTTGCAGCACTGCG 755
 Qy 1546 GCGCTCTGAGATGCTGCTCCGCGGCGCTGACGCAATGGCGGCTTTTCTGCA-GGAC 1604
 Db 756 GCGCTCTGAGATGCTGCTCCGCGGCGCTGACGCAATGGCGGCTTTTCTGCAAGGAG 815
 Qy 1605 CAGTTCGTGAGCACTACGCGCAAGGCGCGCTTCCACCGCGGCGCGGCGCGCGG 1664
 Db 816 CAGTTCGTGAGCACTACGCGCAAGGCGCGCTTCCACCGCGGCGCGGCGCGGCGG 875
 Qy 1665 CCGGCTACGACCGGCGAG 1681
 Db 876 CCGGCGCGGCGAGCCCGG 892

RESULT 9
 LOCUS BM922065 1044 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6707015 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754122

5', mRNA sequence.
 BM922065
 BM922065.1 GI:19372444
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12791 row: a column: 03
 High quality sequence stop: 694.
 Location/Qualifiers
 1..1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5754122"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH-MGC library."

ORIGIN
 Query Match 24.8%; Score 789.8; DB 3; Length 1044;
 Best Local Similarity 97.5%; Pred. No. 7.1e-180;
 Matches 866; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

202 AGTGGCTTGTGCTAGTGAGGGGCTGTGAGATTCTAGTTATGATAATAGACTTGA 261
 |||||
 23 AGTGGCTTGTGCTAGTGAGGGGCTGTGAGATTCTAGTTATGATAATAGACTTGA 82
 |||||
 262 AAAACGAGGAGCGGAGGCGAGTGTTGAGTTGAGACTATGAGCTGAGCACTGCC 321
 |||||
 83 AAAACGAGGAGCGGAGGCGAGTGTTGAGTTGAGACTATGAGCTGAGCACTGCC 142
 |||||
 322 TTTCTCTTATCATGCTCTCTCCAGAAATTTCTCATGTGAGCAGCACTTCCACATT 381
 |||||
 143 TTTCTCTTATCATGCTCTCTCCAGAAATTTCTCATGTGAGCAGCACTTCCACATT 202
 |||||
 382 CAGGCGGATCCGAGCTGTCTGACAGATCCCATTTCTCGACCACTGACCACTGG 441
 |||||
 203 CAGGCGGATCCGAGCTGTCTGACAGATCCCATTTCTCGACCACTGACCACTGG 262
 |||||
 442 AGAGAAGTTCACCTTGAATAAAATTACAGGAGGAGTGAAGATCAGTCACTAGACAGA 501
 |||||
 263 AAGAGAAGTTCACCTTGAATAAAATTACAGGAGGAGTGAAGATCAGTCACTAGACAGA 322
 |||||
 502 ATTTCCTGACTCCAGAGACTCACAGGAGAGCTCCAGAGAAATAAATCTCTCTGCTCT 561
 |||||
 323 ATTTCCTGACTCCAGAGACTCACAGGAGAGCTCCAGAGAAATAAATCTCTCTGCTCT 382
 |||||
 562 TGAGCGGGGTCTCTGAGCTTCGTCACAGTTTGATGCTCTGCTGAGATCCCTACTCTCC 621

Db 383 TGAGCGGGGTCTCTGAGCTTCTGTCAAGTTTCCATGCTGTGTCGACAGATCGTACTCC 442
 |||||
 Qy 622 TCTCTAGTCCAGCCAGCCAGCTCTGCGGCGACCTGCTCCAGTGCATGTTCCGTTACC 681
 |||||
 Db 443 TCTCTAGTCCAGCCAGCCAGCTCTGCGGCGACCTGCTCCAGTGCATGTTCCGTTACC 502
 |||||
 Qy 682 CCGGCGGAGCGAGCGGCGGCGACCCGCTTCTCATCGGAGCCCTAGCGGCTACATGG 741
 |||||
 Db 503 CCGGCGGAGCGAGCGGCGGCGACCCGCTTCTCATCGGAGCCCTAGCGGCTACATGG 562
 |||||
 Qy 742 CCCACACCGGCTCATCCAGACGAGACCTTACACAGCTCTCTGCTCCAGTCTCGCGCG 801
 |||||
 Db 563 CCCACACCGGCTCATCCAGACGAGACCTTACACAGCTCTCTGCTCCAGTCTCGCGCG 622
 |||||
 Qy 802 AGGATATCCCAACCGCGCGCTACCTCTTACCAACGAGTACGAGCTCTCTTACCAAGAG 861
 |||||
 Db 623 AGGATATCCCAACCGCGCGCTACCTCTTACCAACGAGTACGAGCTCTCTTACCAAGAG 682
 |||||
 Qy 862 CTCCGTTCTACCACTTCTCTCCACGAGCGGCGCTGTCGCCGCAAGACAGGTGT 921
 |||||
 Db 683 CTCCGTTCTACCACTTCTCTCCACGAGCGGCGCTGTCGCCGCAAGACAGGTGT 742
 |||||
 Qy 922 ACTGTGCAACAGGCGCCCTTTGAGCTGAATTTGACCG-GCACCAACGAGATGATCAT 979
 |||||
 Db 743 ACTGTGCAACAGGCGCCCTTTGAGCTGAATTTGACCGGACCAACGAGATGATCAT 802
 |||||
 Qy 980 CACCAACA-GGAGAGCGAGTGTTC-TTTTTAAGTTTAAACATTTCTGCTCTC--- 1034
 |||||
 Db 803 CACCAACAAGGAGAGCGAGTGTTCCTTTTAAAGTTTAAACATTTCTGCTCTCG 862
 |||||
 Qy 1035 GATCCACGCGCTCATTAACATTTT--GTGATGATTTGGCGG 1080
 |||||
 Db 863 ATCCCAAGGCTCATTAACATTTTGTGATGATTTGGGG 910
 |||||

RESULT 10
 B1914132 945 bp mRNA linear EST 16-OCT-2001
 LOCUS 603182077F1 NIH-MGC_121 Homo sapiens cDNA clone IMAGE:5246298 5',
 DEFINITION mRNA sequence.
 B1914132
 ACCESSION B1914132.1 GI:16178444
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11621 row: i column: 19
 High quality sequence stop: 711.
 Location/Qualifiers
 1..945
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246298"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

FEATURES
 source

QY 442 AGAGAAGTTCACTTGAAGAAATTTACAGGGGATGACGATGATGATACAGACA 501
 Db 240 AGAGAAGTTCACTTGAAGAAATTTACAGGGGATGACGATGATGATACAGACA 299
 QY 502 ATTTTCCTGACTCCAGAGACTCAGAGGGGAGCTCCAGAGAAAGTAACTCTCTCTCT 561
 Db 300 ATTTTCCTGACTCCAGAGACTCAGAGGGGAGCTCCAGAGAAAGTAACTCTCTCTCT 359
 QY 562 TGGACGGGGTCTCTGAGCTTCTGCTCAGATTTGATGAGCTCTGCTGAGATGCTTCT 621
 Db 360 TGGACGGGGTCTCTGAGCTTCTGCTCAGATTTGATGAGCTCTGCTGAGATGCTTCT 419
 QY 622 TCTCTCAGTCCAGCCAGCCAGCTGAGGGGAGCTGCTCCAGTGCATGTTCCCGTACC 681
 Db 420 TCTCTCAGTCCAGCCAGCCAGCTGAGGGGAGCTGCTCCAGTGCATGTTCCCGTACC 479
 QY 682 CCGGCGCAGCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 741
 Db 480 CCGGCGCAGCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 539
 QY 742 CCGGCGCAGCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 801
 Db 540 CCGGCGCAGCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 599
 QY 802 AGGGAATACCCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 861
 Db 600 AGGGAATACCCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCTCAGTCCAGT 658
 QY 862 CTCGGTTCACAGTCTCTCTCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCT 921
 Db 659 CTCGGTTCACAGTCTCTCTCAGCCAGCCAGCCAGCTGCTCAGTCCAGCCAGCTGCT 718
 QY 922 ACCTGTGCAACGCGCCCTTGGCTGAAATTTCAACGCGGACCAAAAGAGATGATCATCA 981
 Db 719 ACCTGTGCAACGCGCCCTTGGCTGAAATTTCAACGCGGACCAAAAGAGATGATCATCA 776
 QY 982 CCAAAACAGGGAGGGGAGCTGTTCTCTTTTAAAGTTTAAATTTGATGATGATCATCA 1041
 Db 777 CCAAAACAGGGAGGGGAGCTGTTCTCTTTTAAAGTTTAAATTTGATGATGATCATCA 834
 QY 1042 CCGCTCATTAACAATTTTGTGATGATGATTTGGCGGATCCCAATCACTGAGAGTTTC 1101
 Db 835 GGTTCATTAACAATTTTGTGATGATGATTTGGCGGATCCCAATCACTGAGAGTTTC 888
 QY 1102 AAGGAGGCAATAGGTTCTTCTGCGGCAAGCGGACCAATGTCAGGAAATGCGGTTCT 1161
 Db 889 AAGGAGGCAATAGGTTCTTCTGCGGAGG--AAGCGGACCCCAATTGCG--GGGAACCGGTTCT 944
 QY 1162 ATATGATCCGAGTTCCCCCA 1182
 Db 945 ATTTGCGTCCCGTTCCCA 965
 RESULT 12
 LOCUS AM003023/c 831 bp mRNA linear EST 08-MAR-2000
 DEFINITION wg61d04.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2475751.3
 similar to SW:TBRI_HUMAN Q16650 T-BRAIN-1 PROTEIN; contains PTRS.t3
 TARI repetitive element ;, mRNA sequence.
 ACCESSION AM003023
 VERSION AM003023.1 GI:5849861
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 831)
 NCI-GAP <http://www.ncbi.nlm.nih.gov/ncigap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-GAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLN at:
www-bio.llnl.gov/bdip/image/image.html
 Insert Length: 880 Std Error: 0.00
 Seq primer: -40UP from Gldco
 High quality sequence stop: 458.
 location/Qualifiers
 1..831
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2475751"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_id="NCI CGAP GC6"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
 from the normalized library NCI CGAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (cloneIDs
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 23.9%; Score 760; DB 1; Length 831;
 Best Local Similarity 94.8%; Pred. No. 1;e-172;
 Matches 791; Conservative 0; Mismatches 40; Indels 3; Gaps 1;

QY 1146 CAAGGAATCGGCTATATGATCCGGAATCCCAACCTGGGCTCACTGATGCGC 1205
 Db 831 CGAGGAATCGGCTATATGATCCGGAATCCCAACCTGGGCTCACTGAA---GC 775
 QY 1206 CAAGGAATCTTTTGGAAATTTAACTTACGAACAACAAGAGCTTCAATTAACAT 1265
 Db 774 GCAAGGATCTTTTGGAAATTTAACTTACGAACNCAAGAGCTTCAATTAACAAAT 715
 QY 1266 GGGCAGATGGTGTTTTACATCTTGTGCAAGTACAGCCCGGCTGCAATGTGTGAA 1325
 Db 714 GGGCAGATGGTGTGTTCACATCTTGTGCAAGTACAGCNCGCTGCAATGTGTGAA 655
 QY 1326 GTGAACGAGGACGGGACGGAAGCACTAGCCAGCCGGGCGGTGCAAGCGTCACTTTC 1385
 Db 654 GTGANCGAGNACAGACGGAAGCACTAGCCAGCCGAGCGGTGCAAGCGTCACTTTC 595
 QY 1386 CTTGAGACTCAGTTTATCTGCGGCTGACCGCTACCAAGAACAGATATTACAACTGAAA 1445
 Db 594 CTTGAGACTCAGTTTATCTGCGGCTGACCGCTACCAAGAACAGATATTACAACTGAAA 535
 QY 1446 ATAGATCAACAACCTTTTGGCAAAAGATTTGGGATATTATGACAGATTTACACGGC 1505
 Db 534 ATAGATCAACAACCTTTTGGCAAAAGATTTGGGATATTATGACAGATTTACACGGC 475
 QY 1506 TGTGACATGAGACCGGCTGACCGGCTGCGCAAGACTGCGGCTGCGAGATGTCGCC 1565
 Db 474 TGTGACATGAGACCGGCTGACCGGCTGCGCAAGACTGCGGCTGCGAGATGTCGCC 415
 QY 1566 GGGGCGGCTGACCGGCTGACCGGCTGCTTTCTTGTGAGAGCAAGTTCGAGCACTACGCC 1625
 Db 414 GGGGCGGCTGACCGGCTGACCGGCTGCTTTCTTGTGAGAGCAAGTTCGAGCACTACGCC 355
 QY 1626 AAGGCGGCTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1685
 Db 354 AAGGCGGCTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 295

QY 1686 CCGCACCAACAGGCGCTGCTGTGCGCCAGCAGAGCCGAGAGCCCGCCCTTGGCG 1745
 DB 294 CCGCACCAACAGGCGCTGCTGTGCGCCAGCAGAGCCGAGAGCCCGCCCTTGGCG 235
 QY 1746 CAACGCTGTTTGTGAGCGCGGCAACACCGGCTGAGCTTGCGGGGCTCGGCTTATGAC 1805
 DB 234 CAACGCTGTTTGTGAGCGCGGCAACACCGGCTGAGCTTGCGGGGCTCGGCTTATGAC 175
 QY 1806 ACAGCAGCAGGACTTCCGCGGCAACGCGGCAACGCTGCTCTTTCACGCGGCGGCGGCTG 1865
 DB 174 ACAGCAGCAGGACTTCCGCGGCAACGCGGCAACGCTGCTCTTTCACGCGGCGGCGGCTG 115
 QY 1866 AAGCGCTGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCGCTGCGCTTATGAGCCGAC 1925
 DB 114 AAGCGCTGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCGCTGCGCTTATGAGCCGAC 55
 QY 1926 CCGTGGGCTGGGCGCGCGGCAAGCTGCGGCAAGCTGCGGCGGCAAGCTGCGGCG 1979
 DB 54 CCGTGGGCTGGGCGCGCGGCAAGCTGCGGCAAGCTGCGGCGGCAAGCTGCGGCG 1
 RESULT 13
 LOCUS CN643299 910 bp mRNA linear EST 12-MAY-2004
 DEFINITION ILLUMIGEN MCO_7632 Katze_MBR Macaca mulatta cDNA clone 1810W:3780
 5' similar to Bases 1 to 810 highly similar to human TBR1
 (Hs_210862), mRNA sequence.
 ACCESSION CN643299
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopithecoidea; Cercopithecinae; Macaca.
 1 (bases 1 to 910)
 MAGNES,C.T., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
 Prohl,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
 Iadonato,S.P.
 Analysis of the Macaca mulatta transcriptome and the sequence
 divergence between Macaca and human
 Genome Biol. 6 (7), R60 (2005)
 15998449
 JOURNAL Contact: C. Magnes
 PUBMED illuminigen Biosciences Inc.
 COMMENT 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagnes@illuminigen.com
 Sequenced on 2004.01.02. 826 Q20 bases.
 PCR Primers
 FORWARD: CCCTCACTAAGGAGAACAAA
 BACKWARD: CACTATAGGCGCAATTTGGTA
 Insert Length: 910 Std Error: 0.00
 Plate: CL000059 Row: D Column: 08
 Seq primer: CCCTCACTAAGGAGAACAAA
 POLYA=No.
 FEATURES
 source location/Qualifiers
 1..910
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IB10W:3780"
 /sex="female"
 /dev_stage="adult"
 /lab_host="E. coli SOLR"
 /clone_1b="Katze_MBR"
 /note="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;
 Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
 kit (catalog #200400) and ZAP-cDNA Gigapack III Gold
 Cloning Kit (catalog #200450)"

ORIGIN

Query Match 23.5%; Score 749.6; DB 7; Length 910;
 Best Local Similarity 97.0%; Pred. No. 3.8e-170;
 Matches 786; Conservative 0; Mismatches 19; Indels 5; Gaps 2;
 QY 2382 CCGGAGCCCGGAGGCGCCCTCACAGCTCTTCCAGCTCGGCTCCCACTCTCC 2441
 DB 1 CCGGAGCCCGGAGGCGCCCTCACAGCTCTTCCAGCTCGGCTCCCACTCTCC 60
 QY 2442 TTGCGACCCCACTATTATTATGACCTCGATGGCGGTGCGAGCAATTAAGTCAAGT 2501
 DB 61 TTGCGACCCCTCAATTTATTTGACCTCGATGGCGGTGCGAGCAATTAAGTCAAGT 120
 QY 2502 CTCGAGGCTGATTTTAAAC-TTTTTGGACAGAGTCTCTGCAATTAGTCAACGACCT 2560
 DB 121 CTCGAGGCTGATTTTAAACCTTTTTTGCACAGAGTCTCTGCAATTAGTCAACGACCT 180
 QY 2561 TCACTTTGCTGTAACCTTTTGGTTTCTTACTTCTTCTGTTGAGATTAATCTC 2620
 DB 181 TCACTTTGCTGTAACCTTTTGGTTTCTTACTTCTTCTGTTGAGATTAATCTC 240
 QY 2621 CTACATTCCTCCCTCCCTCGTCTTCTCTTAACCTCTAATCTTCTCTTCTTGAATGAA 2680
 DB 241 CTACATTCCTCTTCCCTCGTCTTCTCTTAACCTCTAATCTTCTCTTCTTGAATGAA 300
 QY 2681 CTCTTCACTTTAGAGAGCTGCGGAGTCTGTGAGGAGCGAGCTTCCGACCGGCA 2740
 DB 301 CTCTTCACTTTAGAGAGCTGCGGAGTCTGTGAGGAGCGAGCTTCCGACCGGCA 360
 QY 2741 GTCTGGGCTGCACTTAACGATTAAGATGATGATGATGATGATGATGATGATGATGAT 2800
 DB 361 GTCTGGGCTGCACTTAACGATTAAGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 2801 TCTCTTCTTATCCCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2860
 DB 421 TCTCTTCTTATCCCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 2861 AGTGAACAAAGCTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2920
 DB 481 AGTGAACAAAGCTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 2921 GTTTTATCTGCTGAAGCAAAACAGTCAATGATGATGATGATGATGATGATGATGAT 2980
 DB 541 GTTTTATCTGCTGAAGCAAAACAGTCAATGATGATGATGATGATGATGATGATGAT 600
 QY 2981 CTTTGTCTTCTCTCTCTCTC-ATACCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 3036
 DB 601 CTTTGTCTTCT 660
 QY 3037 GTGAGATTAATTTGAAGGCTCTGAAGCAATGAATGATGATGATGATGATGATGATGAT 3096
 DB 661 GTGAGATTAATTTGAAGGCTCTGAAGCAATGAATGATGATGATGATGATGATGATGAT 720
 QY 3097 ACTTCTCTCAATCCGTTGAGGAAATATTTCTATGTCCTTAATGACAAATAGCT 3156
 DB 721 ACTTCTCTCAATCCGTTGAGGAAATATTTCTATGTCCTTAATGACAAATAGCT 780
 QY 3157 AAGAGAAATCAACCAACCTTTAAAG 3186
 DB 781 AAGAGAAATCAACCAACCTTTAAAG 810
 RESULT 14
 LOCUS DR001136 768 bp mRNA linear EST 17-MAY-2005
 DEFINITION TC115971 Human fetal brain, large insert, pCMV expression library
 Homo sapiens cDNA clone TC115971 5' similar to Homo sapiens T-box,
 brain, 1 (TBR1), mRNA sequence.
 ACCESSION DR001136
 VERSION DR001136.1 GI:66261009
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM	TITLE	JOURNAL	COMMENT
Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homniidae; Homo. 1 (bases 1 to 768)	Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L., Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M., Zhang,X., Jay,G., and He,W. High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts Unpublished (2005) Contract: Kovacs, KP High throughput cDNA Cloning Origene Technologies, Inc. (www.origene.com) 6 Taft Court, Suite 100, Rockville, MD 20850, USA Tel: 301 340 3188 Fax: 301 340 8606 Email: CDNA@origene.com	This EST submission is part of an on-going human full-length cloning project at Origene Technologies, Inc. Please contact Origene for access. Origene Technologies, Inc. 6 Taft Ct. Suite 100 Rockville, MD 20850 Tel: (301) 340-3188 http://www.origene.com Seq primer: PCMV6 5prime forward vector primer, Origene Technologies Inc.	
FEATURES	location/Qualifiers		
source	1..768		
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="TC115971"			
/tissue.type="Fetal Brain"			
/clone.lib="Human fetal brain, large insert, PCMV expression library"			
/note="Organ: Fetal Brain; Vector: pCMV6-XL4; Site.1: ECORI; Site.2: XhoI/SalI compatible end ligatio; Oligo-dT primed reverse transcription optimized for large and GC rich mRNA transcripts. cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"			
ORIGIN			
Query Match	22.7%;	Score 722.4;	DB 8; Length 768;
Best Local Similarity	97.7%;	Pred. No. 1.4e-163;	
Matches	732; Conservative	0; Mismatches	17; Indels 0; Gaps 0
Db	203 GTGCTTTCTGTAGTAGAGGGGCTGTGTGATTTCTAGTTTAATAAGACTTTAA	262	
OY			
Db	9 GGCTTTCTGTCTAGTAGAGGGGCTGTGTGATTTCTAGTTTAATAAGACTTTAA	68	
OY	263 AAACCAAGGACCGGAGGGCGAGTGTTCCAGGTTCTAGAGCTTAGCAGCTGGAGACTGCCT	322	
Db	69 AAACCAAGGACCGGAGGGCGAGTGTTCCAGGTTCTAGAGCTTAGCAGCTGGAGACTGCCT	128	
OY	323 TTCTCTCTTATCATGSCCTCCAGAAGAAATTTCGCATATGAGAGAGAGCTACCACATTC	382	
Db	129 TTCTCTCTTATCATGSCCTCCAGAAGAAATTTCGCATATGAGAGAGAGCTACCACATTC	188	
OY	383 AAGCGGATCCGAGCTGTCTTGAGCAGCATGCCATTTATCTGCAGCACGACACAACCTGGA	442	
Db	189 AAGCGGATCCGAGCTGTCTTGAGCAGCATGCCATTTATCTGCAGCACGACACAACCTGGA	248	
OY	443 GAGAACTTCACTTTGAAAAAATTACAGGGGGATGACGAATCACTCAGATACAGACAA	502	
Db	249 GAGAACTTCACTTTGAAAAAATTACAGGGGGATGACGAATCACTCAGATACAGACAA	308	
OY	503 TTTTCTGACTCCAGGACTCACACAGGGGACGTCCAGAGAAGTAACTCTCTCTGTCTT	562	
Db	309 TTTTCTGACTCCAGGACTCACACAGGGGACGTCCAGAGAAGTAACTCTCTCTGTCTT	368	

OY	563	GGAGGGGCTCTCNAAGTTTGGTCACAATTTCGATGGCTCTGCTGCAGATGCTAACCTCCCT	622
Db	369	GGAGGGGAGTCTCTGAGCTTGGTCACAATTTTCGATGGCTCTGCTGCAGATGCTAACCTCCCT	428
OY	623	CTTCAGTCCAGCGACGACAGTGTGGGGCCACTGCTCCAGTGCATGTTTCCGTAACC	682
Db	429	CTCTCAGTCCAGCCAGCCACAGTGTGGGGCCACTGCTCCAGTGCATGTTTCCGTAACC	488
OY	683	CGGCGAGCAGGAGCCGGGCGACCCCGCTTCTCCATGGGACGACCTTGCCGCTACATGCG	742
Db	489	CGGCGAGCAGGAGCCGGGCGACCCCGCTTCTCCATGGGAGCCCTTGCCGCTACATGCG	548
OY	743	CCACACCCCGGTCAATCAACAAACGAGGCTTCAACAGGCTCTGTCGAATCTTGCGCCGA	802
Db	549	CCACACCCCGGTCAATCAACAAACGAGGCTTCAACAGGCTCTGTCGAATCTTGCGCCGA	608
OY	803	GGGATACCCCAACGAGCCGGCTAACCCCTTACCCACAGACAGTACGGGCCACTCTTACCAAGAGC	862
Db	609	GGGATACCCCAACGAGCCGGCTAACCCCTTACCCACAGACAGTACGGGCCACTCTTACCAAGAGC	668
OY	863	TTCGTTTCAACAGTTTCTCTCCACCAGCCGGGGCTGTGCCCCGGCAACACAGAGTGTA	922
Db	669	TTCGTTTCAACAGTTTCTCTCCACCAGCCGGGGCTGTGCCCCGGCAACAGTGTGA	728
OY	923	CCTGNGACAAGGCCCTTTGGCTGAAT	951
Db	729	CCTGTGACAGAGCCCTTTGGGCTGAAT	757
RESULT 15			
LOCUS	CN805179	910 bp	mRNA linear EST 26-MAY-2004
DEFINITION	ILLUMIGEN MCO 36121 Katze MM8R Macaca mulatta cDNA clone		
	IBUW:13616.5	similar to Baes 1 to 810 highly similar to human	
	TBR1 (Hs_210862), mRNA sequence.		
ACCESSION	CN805179		
VERSION	CN805179.1	GI:47701155	
KEYWORDS	EST.		
SOURCE	Macaca mulatta (rhesus monkey)		
ORGANISM	Macaca mulatta		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Cercopithecoidea; Cercopithecinae; Macaca.		
REFERENCE	1 (bases 1 to 910)		
AUTHORS	Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,		
	Prohl,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and		
	Iadonato,S.P.		
	Analysis of the Macaca mulatta transcriptome and the sequence		
	divergence between Macaca and human		
	Genome Biol. 6 (7), R60 (2005)		
	15998449		
TITLE	Contact: C. Magness		
JOURNAL	Illumigen Biosciences Inc.		
PUBMED	2203 Airport Way S, Suite 450, Seattle, WA 98134, USA		
COMMENT	Tel: 2063780400		
	Fax: 2063780408		
	Email: cmagness@illumigen.com		
	Sequenced on 2004.05.11. 690 Q20 bases.		
	PCR Primers		
	FORWARD: CCTTCACCTAAAGGACAAAA		
	BACKWARD: CACTATAGGCGAATTGGGTA		
	Insert Length: 910 Std Error: 0.00		
	Plate: C1000250 row: F column: 03		
	Seq primer: CCTTCACCTAAAGGACAAAA		
	PolyA=No.		
FEATURES	Location/Qualifiers		
source	1..910		
	/organism="Macaca mulatta"		
	/mol_type="mRNA"		
	/strain="Indian"		
	/db_xref="taxon:9544"		
	/clone="IBUW:13616"		

```

/sex="female"
/dev_frame="adult"
/lab_host="E. coli SOLR"
/clone_id="Katze_MMBR"
/notes="Organ: brain; Vector: Uni-ZAP XR; Site:1: EcoR I;
Site 2: Xho I; Created from Stratiogene ZAP-cDNA Synthesis
kit (catalog #200400) and ZAP-cDNA Gigapack III Gold
Cloning Kit (catalog #200450)"

```

ORIGIN

Query Match	Score	DB 7;	Length
22.68;	720;	910;	

Best Local Similarity	55.0%	freq. no.	5,76	100%			
Matches 774; Conservative	0;	Mismatches	30;	Indels	6;	Gaps	3;

QY	2382	CCCGGACCCCGACGACGCCCTCAACACTCTCTCCCGAGCTCCGCTCCCGCAACTCCTCC	2441
Db	2	CCGTTACCCCGACGACGCCCTCAACAGCTCTTCCCGAGCTCCGCAACCCTACTCTCC	61
QY	2442	TTGGGCAACCCACTATTTTATTGACCCTCGATGGCCGTCGTGACGCAATTAAGTGACGT	2501
Db	62	TTGGGCAACCCCTCATTTTATTGACCCTCGATGGCCGTCGTGACGCAATTAAGTGACGT	121
QY	2502	CTCCGAGCGTGATTTTAAAC-TTTTGTGCAGAGAGTCTGTGCATTTAGCTCACCCGACT	2560
Db	122	CTCCGAGCGTGATTTTAAACCTTTTGTGCAGAGAGTCTGTGCATTTAGGCAACCCGACT	181
QY	2561	TCAACTTTGGCTGAACCTTTGGTTTTCCACTACTCTCTCTGTGAGATTAATCCTC	2620
Db	182	TCAACTTTGGCTGAACCTTTTGGTTTTCCACTACTCTCTCTGTGAGATTAATCCTC	241
QY	2621	CTACAAATTCCTCCCTCCCTCGTCTTCTCTTACCTCCTACTCTCTTCTTTGTAATGAA	2680
Db	242	CTACAAATCTCTTCCCTCCCTCGTCTTCTCTTACCTCCTACTCTCTTCTTTGTAATGAA	301
QY	2681	CTCTTCACTTTAAGAGACTTGGGCACTCTGTGACGAGAGAGAGATTCGACCCGCGCA	2740
Db	302	CTCTTCACTTTAAGAGACTTGGGCACTCTGTGACGAGAGAGATTCGACCCGCGCA	361
QY	2741	GTCCTGGGCTCCACATTACATAGAGATGTGACTTGAACCTGGAACCCAGCGCG	2800
Db	362	GTCCTGGGCTCCACATTACATAGAGATGTGACTTGAACCTGGAACCCAGCGCG	421
QY	2801	TCCTTTCTTATCCCGAGTGAGATGAGATGAGATGATGATGATGATGATGATGATGATGAT	2860
Db	422	TCCTTTCTTATCCCGAGTGAGATGAGATGAGATGAGATGAGATGAGATGATGATGATGAT	481
QY	2861	AGTGAACAACAGCCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2920
Db	482	AGTGAACAACAGCCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	541
QY	2921	GTTTATTAATTCGTCAAGGCAACAAACAGTTCATGCTTAACTTTTTCCTTTCCTT	2980
Db	542	GTTTATTAATTCGTCAAGGCAACAAACAGTTCATGCTTAACTTTTTCCTTTCCTT	601
QY	2981	CTTTCCTTTCTTTCTCTCTCTC---ATACTTCTCTTCTCTCTCTTTTAATTTCTT	3036
Db	602	CTTTCCTTTCTTTCTCTCTCTCTCAATCTTCTCTCTCTCTCTCTTTTAATTTCTT	661
QY	3037	GCGAATATATTTCTTAAGAGGCTTAAGAAACATGAATACAGTACAGTACAGTACAGT	3096
Db	662	GCGAATATATTTCTTAAGAGGCTTAAGAAACATGAATACAGTACAGTACAGTACAGT	721
QY	3097	ACTTCTCCTCAATCCGTTGCATGAATAATTAATTAATGATGATGATGATGATGATGATGAT	3156
Db	722	ACTTCTCCTCAATCCGTTGCATG-ATAATTAATTAATGATGATGATGATGATGATGATGAT	780
QY	3157	AAGGAGATTCACCCCAACCTTTTAAAG 3186	
Db	781	AAGGAGATTCACCCCAACCTTTTAAAG 810	

Search completed: December 20, 2005, 15:56:22
Job time : 12035 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 20, 2005, 07:33:43, Search time 543 Seconds
(without alignments)
10429.671 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 GAGGTGATCCTCCTAATAA.....CACCAACACCTTAAAG 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:
1: /cgn2_6/prodata/1/ina/1_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/H_COMB.seq:*
6: /cgn2_6/prodata/1/ina/BCTUS_COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/prodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2894	90.8	2912	3	US-09-949-016-4834
2	1695.4	53.2	12681	3	US-09-949-016-16576
3	514	16.1	514	3	US-09-775-398-37
4	384.6	12.1	388	3	US-09-513-999C-885
5	374.6	11.8	1704	3	US-09-188-811-1
6	374.6	11.8	2494	3	US-09-189-760-1
7	374.6	11.8	2494	3	US-09-514-422-1
8	316.2	9.9	2575	3	US-09-949-016-3434
9	316.2	9.9	2589	3	US-09-949-016-976
10	235	7.4	2253	3	US-09-513-999C-22243
11	150.6	4.7	1798	3	US-09-949-016-656
12	150.6	4.7	1799	3	US-09-949-016-2908
13	146	4.6	2433	3	US-09-949-016-4821
14	146	4.6	2433	3	US-09-949-016-4822
15	146	4.6	3113	3	US-10-028-272-1
16	144.4	4.5	2470	3	US-09-949-016-1061
17	144.4	4.5	2472	3	US-09-949-016-2225
18	142.8	4.5	2454	3	US-09-949-016-4491
19	127.6	4.0	2441	3	US-09-949-016-29
20	116.8	3.7	2279	3	US-09-949-016-4013
21	116.6	3.7	6717	3	US-10-082-272-1
22	113.2	3.6	1521	3	US-09-949-016-5766
23	111.2	3.5	2238	3	US-08-939-366-5
24	105.4	3.3	1881	3	US-09-774-528-52

25	105.4	3.3	1881	3	US-10-120-988-52	Sequence 52, Appl
26	103.8	3.3	3179	3	US-09-949-016-1040	Sequence 1040, Ap
27	97.6	3.1	2822	3	US-09-949-016-5104	Sequence 5104, Ap
28	97.6	3.1	2882	3	US-09-949-016-724	Sequence 724, App
29	96.8	3.0	3940	3	US-09-799-451-919	Sequence 919, App
30	88	2.8	632	3	US-09-949-016-1494	Sequence 1494, Ap
31	86.4	2.7	7218	3	US-08-232-463-14	Sequence 14, Appl
32	84.2	2.6	1143	3	US-08-774-528-60	Sequence 60, Appl
33	84.2	2.6	1143	3	US-10-120-988-60	Sequence 60, Appl
34	82.8	2.6	16871	3	US-09-949-016-12718	Sequence 12718, A
35	82.8	2.6	16874	3	US-09-949-016-15176	Sequence 15176, A
36	77	2.4	2516	3	US-10-104-047-485	Sequence 485, App
37	77	2.4	9454	3	US-09-949-016-12398	Sequence 12398, A
38	77	2.4	9455	3	US-09-949-016-14650	Sequence 14650, A
39	72.6	2.3	77536	3	US-09-410-5518-1	Sequence 1, Appl
40	72.6	2.3	77536	3	US-09-940-316B-1	Sequence 1, Appl
41	72.2	2.3	601	3	US-09-949-016-38383	Sequence 38383, A
42	72.2	2.3	601	3	US-09-949-016-170650	Sequence 170650, A
43	72.2	2.3	601	3	US-09-949-016-170658	Sequence 170658, A
44	72.2	2.3	8438	2	US-07-945-283-1	Sequence 1, Appl
45	72.2	2.3	15621	3	US-09-949-016-16563	Sequence 16563, A

ALIGNMENTS

RESULT 1
US-09-949-016-4834
Sequence 4834, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSeq for Windows Version 4.0
SEQ ID NO 4834
LENGTH: 2912
TYPE: DNA
ORGANISM: Human
US-09-949-016-4834
Query Match 90.8%; Score 2894; DB 3; Length 2912;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2908; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
279 GCGAGGTGTCAGGTTCTGAGCTTTCAGCTGAGCACTGCTTCTTCTTCTGAG 338
1 GCGAGGTGTCAGGTTCTGAGCTTTCAGCTGAGCACTGCTTCTTCTTCTGAG 60
339 CTCCTCAGAAATTTCTCAATGATGAGCAGCAGCTTCCATTCAGGCGATTCGAGCTT 398
61 CTCCTCAGAAATTTCTCAATGATGAGCAGCAGCTTCCATTCAGGCGATTCGAGCTT 120
399 GTCCTGACAGATTCATCCATTAATCTTCGACCACTGACCACTTCGAGGAGTTCACCTTTG 458
121 GTCCTGACAGATTCATCCATTAATCTTCGACCACTGACCACTTCGAGGAGTTCACCTTTG 180
459 AAAAAAATTCAGAGGGGATGAGCAATGATGAGCAATTCCTGACTTCAG 518
181 AAAAAAATTCAGAGGGGATGAGCAATGATGAGCAATTCCTGACTTCAG 240
519 GACTCAGCAGGAGGAGTTCAGAGAGTAATCTCTCTCTGAGCGGGTCTCTGAG 578

Db 241 GACTCACAGGAGAGTCCAGAGAAATACTCTCTGTTGTGAGCGGGGCTCTGAG 300
QY 579 CTTGCTACAGTTTGTGATGAGCTCTGTGAGATTCCTTACCTCTCTTCACTGTCAGCCAG 638
Db 301 CTTTCGTACAGTTTGTGATGAGCTCTGTGAGATTCCTTACCTCTCTTCACTGTCAGCCAG 360
QY 639 CCACAGTCTGCGGCGACTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAGGACCG 698
Db 361 CCACAGTCTGCGGCGACTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAGGACCG 420
QY 699 GCGCACCCCGCTTCTCCATGCGCAGCCCTAGCCGCTCATGAGCCACACCCGGTCACTC 758
Db 421 GCGCACCCCGCTTCTCCATGCGCAGCCCTAGCCGCTCATGAGCCACACCCGGTCACTC 480
QY 759 ACCAAGGAGGCTTACAAAGGCTCTCTGTCACATCTCGCGCAGGAGTACCCACGCGCC 818
Db 481 ACCAAGGAGGCTTACAAAGGCTCTCTGTCACATCTCTCGCGCAGGAGTACCCACGCGCC 540
QY 819 GAGTACCCCTTACCAAGCAGATACGAGCCTCTTACCAAGAGTCCGTTCTACCAAGTTC 878
Db 541 GAGTACCCCTTACCAAGCAGATACGAGCCTCTTACCAAGAGTCCGTTCTACCAAGTTC 600
QY 879 TCCCTCACCCAGCCGCGGCTGTGTGCTCCGCAAGCAAGTGTACCTGTGCAACAGGCC 938
Db 601 TCCCTCACCCAGCCGCGGCTGTGTGCTCCGCAAGCAAGTGTACCTGTGCAACAGGCC 660
QY 939 CTTTGGCTGAATTTTCAACCGGCAACCAAGAGATGATATCAACCAACAGGGAAGGGCG 998
Db 661 CTTTGGCTGAATTTTCAACCGGCAACCAAGAGATGATATCAACCAACAGGGAAGGGCG 720
QY 999 ATGTTTCCCTTTTAAAGTTTAAATTTCTGTCTGATCCCAAGGCTCATTAACAATTT 1058
Db 721 ATGTTTCCCTTTTAAAGTTTAAATTTCTGTCTGATCCCAAGGCTCATTAACAATTT 780
QY 1059 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAAGTGGTT 1118
Db 781 TTTTGTGATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAAGTGGTT 840
QY 1119 CCTTGGCGGAAGCGGACCAATGTGCAAG---GAAATCGGGTCTAATTCATCCGGA 1174
Db 841 CCTTGGCGGAAGCGGACCAATGTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 900
QY 1175 TTTCCGCCAACACTGGGGGCTCACTGATGCGCCAAAGAAATCTTTTGGAAAATTTAAACT 1234
Db 901 TTTCCGCCAACACTGGGGGCTCACTGATGCGCCAAAGAAATCTTTTGGAAAATTTAAACT 960
QY 1235 TACGAACAACAAGAGCTTCAATTAACAATGAGGAGATGTGTGTTTACAGTCTTGCA 1294
Db 961 TACGAACAACAAGAGCTTCAATTAACAATGAGGAGATGTGTGTTTACAGTCTTGCA 1020
QY 1295 CAAGTACCAAGCCCCCGCTGATGTGTGAAAGTGAACGAGACGCGCAGGAGACACTAG 1354
Db 1021 CAAGTACCAAGCCCCCGCTGATGTGTGAAAGTGAACGAGACGCGCAGGAGACACTAG 1080
QY 1355 CCAGCCCGGCGCGTGCAGACGTTCACTTCCCTGAGACTAGTTCATGCGCTCACGC 1414
Db 1081 CCAGCCCGGCGCGTGCAGACGTTCACTTCCCTGAGACTAGTTCATGCGCTCACGC 1140
QY 1415 CTACCAGAACACGAGATATTACAACAATAAGATACAAACCTTTTGGAAAAGATT 1474
Db 1141 CTACCAGAACACGAGATATTACAACAATAAGATACAAACCTTTTGGAAAAGATT 1200
QY 1475 TGGGATTAATTATGACAGATCTACACCGGCTGTGACATGACCGGCTTGACCCCTCGCG 1534
Db 1201 TGGGATTAATTATGACAGATCTACACCGGCTGTGACATGACCGGCTTGACCCCTCGCG 1260
QY 1535 CAACGACTGCGCGCTGCAATGTGTGCGCGGGGCGGCTAGCGCATGCGCGGCTTT 1594
Db 1261 CAACGACTGCGCGCTGCAATGTGTGCGCGGGGCGGCTAGCGCATGCGCGGCTTT 1320
QY 1595 CCTGACGACCAAGTTCTGTAGGACACTACGCAAGGCGCGCTTCAACCCGCGCGCGCGCG 1654
Db 1321 CCTGACGACCAAGTTCTGTAGGACACTACGCAAGGCGCGCTTCAACCCGCGCGCGCGCG 1380

QY 1655 GGGCCCGGCGCGGATACGAGCAGAGTGCAGACACCAACGAGGCTGTGCGCGCA 1714
Db 1381 GGGCCCGGCGCGGATACGAGCAGAGCAGAGCAGAGCAGACCAACGAGGCTGTGCGCGCA 1440
QY 1715 GCAAGCCGAGAGACCCGCGCGCGCTTCCGCGCAACGCTGTGTTTGAACCGCGCGCAACAA 1774
Db 1441 GCAAGCCGAGAGACCCGCGCGCGCTTCCGCGCAACGCTGTGTTTGAACCGCGCGCAACAA 1500
QY 1775 CCGGCTGAGACTTCCGCGCGCTTCCGCGCTATGACACCGGCCAACGACTTCCGCGCGCAAC 1834
Db 1501 CCGGCTGAGACTTCCGCGCGCTTCCGCGCTATGACACCGGCCAACGACTTCCGCGCGCAAC 1560
QY 1835 CAAGCTGCTCTTACGCGCGCGCGGAGTGAAGGCGCTGCGTGCAGAGCTGCAGGCTG 1894
Db 1561 CAAGCTGCTCTTACGCGCGCGCGGAGTGAAGGCGCTGCGTGCAGAGCTGCAGGCTG 1620
QY 1895 CACTGGCGCGCGCTGCGCTTACGCGCGACCGTGTGAGGCTGAGGCGCGCGAGTCCGCC 1954
Db 1621 CACTGGCGCGCGCTGCGCTTACGCGCGACCGTGTGAGGCTGAGGCGCGCGAGTCCGCC 1680
QY 1955 GCAATATGCGCGGACCAAGTGTGAGGCTGTGCTGCTGCTGCGCGCAACGCGCGCGCG 2014
Db 1681 GCAATATGCGCGGACCAAGTGTGAGGCTGTGCTGCTGCTGCGCGCAACGCGCGCGCG 1740
QY 2015 CGCGCGCGCATGGCGCGCGCGCAATCCCTACCTGAGGAGGAGGCGAGGCGCTGCGCGCG 2074
Db 1741 CGCGCGCGCATGGCGCGCGCGCAATCCCTACCTGAGGAGGAGGCGAGGCGCTGCGCGCG 1800
QY 2075 CGAGCGCTCGCGCTGCG 2134
Db 1801 CGAGCGCTCGCGCTGCG 1860
QY 2135 TTTCAAGCTGATTCGAGAGCGCTTCTCGATTAATTCATGATGACTCCAGGACTGCGGAT 2194
Db 1861 TTTCAAGCTGATTCGAGAGCGCTTCTCGATTAATTCATGATGACTCCAGGACTGCGGAT 1920
QY 2195 TTTAGAGAGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2254
Db 1921 TTTAGAGAGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
QY 2255 GTCCCGCTCAAGAGCGAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2314
Db 1981 GTCCCGCTCAAGAGCGAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2040
QY 2315 CATTAGCGGCTAATATGCTTCTACTGCGACAGTAAAGGCGCGCGCGCGCGCGCGCGCG 2374
Db 2041 CATTAGCGGCTAATATGCTTCTACTGCGACAGCTAAGCGCGCGCGCGCGCGCGCGCG 2100
QY 2375 GCGCGCGCGCGAGCCCGCGAGCGAGCCCTCAAGCTCTTCCCGAGCTCCGCGCGCGCA 2434
Db 2101 GCGCGCGCGCGAGCCCGCGAGCGAGCCCTCAAGCTCTTCCCGAGCTCCGCGCGCGCA 2160
QY 2435 CTCTCTCTTGGCAACCACTGATTTATTTGACCTGTGATGCGGTGCGAGCAATAG 2494
Db 2161 CTCTCTCTTGGCAACCACTGATTTATTTGACCTGTGATGCGGTGCGAGCAATAG 2220
QY 2495 TGCAGGTCTCGAGCGGATTTTAACTTTTTCAGACAGAGTGTGCAATTAGCTCAC 2554
Db 2221 TGCAGGTCTCGAGCGGATTTTAACTTTTTCAGACAGAGTGTGCAATTAGCTCAC 2280
QY 2555 CGACCTTCAACTTTGCTGTAACCTTTTGTGTTTCTACTTACTCTTCTTGTGAGTT 2614
Db 2281 CGACCTTCAACTTTGCTGTAACCTTTTGTGTTTCTACTTACTCTTCTTGTGAGTT 2340
QY 2615 ATCTCTTACAAATTCCTTCCCTGCTGCTTCTTACCTGCTACTTCTTCTTCTGTA 2674
Db 2341 ATCTCTTACAAATTCCTTCCCTGCTGCTTCTTACCTGCTACTTCTTCTTCTGTA 2400
QY 2675 ATGAAGCTTTCACCTTTAGAGACCTGAGGAGTCTGTGAGGACGAGGAGGAGGAGGAGG 2734
Db 2401 ATGAAGCTTTCACCTTTAGAGACCTGAGGAGTCTGTGAGGAGGAGGAGGAGGAGGAGG 2460

QY 2735 GCCAAGTCTGGCTCCACATTAACCATAGATTTGACTTGAACCTGACCCACC 2794
 DB 2461 CGCCAAAGCTCGGCTCCACATTAACCATAGATTTGACTTGAACCTGACCCACC 2520
 QY 2795 AGCGGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGATGAT 2854
 DB 2521 AGCGGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGATGAT 2580
 QY 2855 AATTTAGTGAACAAGCCTGTGAATGATGATGATGATGATGATGATGATGATGAT 2914
 DB 2581 AATTTAGTGAACAAGCCTGTGAATGATGATGATGATGATGATGATGATGATGAT 2640
 QY 2915 TGGCTAGTTTATTTCTGTCAAGCACAACCAAGTTGATGATGATGATGATGAT 2974
 DB 2641 TGGCTAGTTTATTTCTGTCAAGCACAACCAAGTTGATGATGATGATGATGAT 2700
 QY 2975 TCCCTTCTTGGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3034
 DB 2701 TCCCTTCTTGGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 QY 3035 TTGTAGATTAATTTCTAAGAGCTCTAAGAACATGAATACTCAGTATGATGAT 3094
 DB 2761 TTGTAGATTAATTTCTAAGAGCTCTAAGAACATGAATACTCAGTATGATGAT 2820
 QY 3095 CCAGTTCTCTCAATCCGTTGATGATGAATTAATTAATTAATTAATTAATTAAT 3154
 DB 2821 CCAGTTCTCTCAATCCGTTGATGATGAATTAATTAATTAATTAATTAATTAAT 2880
 QY 3155 CTAAAGAGATCCACCAACCACTTTAAG 3186
 DB 2881 CTAAAGAGATCCACCAACCACTTTAAG 2912

RESULT 2

US-09-949-016-16576
 / Sequence 16576, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 16576
 / LENGTH: 12681
 / TYPE: DNA
 / ORGANISM: Human
 / US-09-949-016-16576

Query Match 53.2%; Score 1695.4; DB 3; Length 12681;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1490 CACGATCTACACCGGCTGTGACATGACGCGCTGACCCCTCCGCCAACGACTGCGCG 1549
 DB 8985 CAGAGATTACACCGGCTGTGACATGACGCGCTGACCCCTCCGCCAACGACTGCGCG 9044
 QY 1550 CTCGACAGATCGTCCGCGGCGCGCTAAGCATGCGCGCTCTTTCTTCTGACGACGAT 1609
 DB 9045 CTCGACAGATCGTCCGCGGCGCGCTAAGCATGCGCGCTCTTTCTTCTGACGACGAT 9104
 QY 1610 CGTGAGACACTACCGCAAGCGCTTCCACCCGCGGCGCGCGCGCGCGCGCGCG 1669
 DB 9105 CGTGAGACACTACCGCAAGCGCTTCCACCCGCGGCGCGCGCGCGCGCGCGCGCG 9164

QY 1670 TACGACCGAGCGGTGCGGACACCAAGGCTGCTGTCCGCGAGGAGCGGAGACCC 1729
 DB 9165 TACGACCGAGCGGTGCGGACACCAAGGCTGCTGTCCGCGAGGAGCGGAGACCC 9224
 QY 1730 GGGGCGCCCTCGCGGACGCTGTGTTGATCGCGCGGACCAACCGGCTGTGACTTCCG 1789
 DB 9225 GGGGCGCCCTCGCGGACGCTGTGTTGATCGCGCGGACCAACCGGCTGTGACTTCCG 9284
 QY 1790 GGCTCGGCTTANAGACAGCGGCACTTCCGCGGACGCGGCGGACGCTGCTCTTA 1849
 DB 9285 GGCTCGGCTTANAGACAGCGGCACTTCCGCGGACGCGGCGGACGCTGCTCTTA 9344
 QY 1850 CGCGGCGGCGGCTGGAAGCGCTGCGCTGACAGGCTGACGCTGACGCTGCGCGCT 1909
 DB 9345 CGCGGCGGCGGCTGGAAGCGCTGCGCTGACAGGCTGACGCTGACGCTGCGCGCT 9404
 QY 1910 CGGCTACTACGCGGACCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG 1969
 DB 9405 CGGCTACTACGCGGACCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9464
 QY 1970 CAAGTGGGCTGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2029
 DB 9465 CAAGTGGGCTGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9524
 QY 2030 CGGCGGCAATGCTTACCTGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCT 2089
 DB 9525 CGGCGGCAATGCTTACCTGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCT 9584
 QY 2090 GCGCGCGGCGGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2149
 DB 9585 GCGCGCGGCGGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 9644
 QY 2150 GACGCTCTCTGATCAAGTCAATGATCTCCAGGAGCTTCCGCGGAGTTTACGAGGCGCA 2209
 DB 9645 GACGCTCTCTGATCAAGTCAATGATCTCCAGGAGCTTCCGCGGAGTTTACGAGGCGCA 9704
 QY 2210 GCGGAGGCGGATCTCGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGAG 2269
 DB 9705 GCGGAGGCGGATCTCGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGAG 9764
 QY 2270 CGAGGTGCTGCGCCAGCGGAGCTGCGGAGAGGAGTGCAGGAGCTTAAAGCGGCTACTA 2329
 DB 9765 CGAGGTGCTGCGCCAGCGGAGCTGCGGAGAGGAGTGCAGGAGCTTAAAGCGGCTACTA 9824
 QY 2330 TGGCTTCTACTGCAAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 2389
 DB 9825 TGGCTTCTACTGCAAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 9884
 QY 2390 CCCAGCGGCGGCTGACAGCTCTTCCGAGCTCCGCGCTCCGAGACTCTCTTCCGAGC 2449
 DB 9885 CCCAGCGGCGGCTGACAGCTCTTCCGAGCTCCGCGCTCCGAGACTCTCTTCCGAGC 9944
 QY 2450 CCAGTCAATTTATTTGACCGCTGATGCGCTGCGAGGAGATTAAGTGAAGTCTCGAGC 2509
 DB 9945 CCAGTCAATTTATTTGACCGCTGATGCGCTGCGAGGAGATTAAGTGAAGTCTCGAGC 10004
 QY 2510 GTGATTTTAACTTTTGTGACAGAGCTCTGCAATTAAGTCAACGACTTCAACTTTG 2569
 DB 10005 GTGATTTTAACTTTTGTGACAGAGCTCTGCAATTAAGTCAACGACTTCAACTTTG 10064
 QY 2570 CTGTAAACTTTTGTGTTTCTTCAATTAAGTCTCTGCGGAGTTATCCCTCAAAATTC 2629
 DB 10065 CTGTAAACTTTTGTGTTTCTTCAATTAAGTCTCTGCGGAGTTATCCCTCAAAATTC 10124
 QY 2630 CCTGCGGCTGCTTCTCTTACTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2689
 DB 10125 CCTGCGGCTGCTTCTCTTACTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 10184
 QY 2690 TTTAGAGACCTGGGAGATGCTGCTGTCAGGAGAGGATTCGAGCCGCAAGTCTCGGCG 2749
 DB 10185 TTTAGAGACCTGGGAGATGCTGCTGTCAGGAGAGGATTCGAGCCGCAAGTCTCGGCG 10244

QY 2750 TCCACATTAACCATGAGATGTTGACTCTAGAACCTGAGACCCAGCCGCGCTCTTTCTT 2809
| | | | |
Db 10245 TCCACATTAACCATGAGATGTTGACTCTAGAACCTGAGACCCAGCCGCGCTCTTTCTT 10304
| | | | |
QY 2810 ATCCCCGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2869
| | | | |
Db 10305 ATCCCCGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 10364
| | | | |
QY 2870 AAGCTGTGAATGATGTGACATAGTGTAAATTTATGTGAAGAAATGGCTATGTTTATTT 2929
| | | | |
Db 10365 AAGCTGTGAATGATGTGACATAGTGTAAATTTATGTGAAGAAATGGCTATGTTTATTT 10424
| | | | |
QY 2930 CTCGTCAAGGACCAAAACAGGTTCAATGCTTAACCTTTTCTTCTTCTTCTTCTTCTT 2989
| | | | |
Db 10425 CTCGTCAAGGACCAAAACAGGTTCAATGCTTAACCTTTTCTTCTTCTTCTTCTTCTT 10484
| | | | |
QY 2990 TCTTTCT 3049
| | | | |
Db 10485 TCTTTCT 10544
| | | | |
QY 3050 CTAAAGAGGCTTAGAAACATGAATACAGTAGAGTAGAGTTTCCCACTTCTCTCAAT 3109
| | | | |
Db 10545 CTAAAGAGGCTTAGAAACATGAATACAGTAGAGTAGAGTTTCCCACTTCTCTCAAT 10604
| | | | |
QY 3110 CCGTTCATGAATATTAATTAATGAGCTTAATGACACAAATAGCTAAGAGAAATCCAC 3169
| | | | |
Db 10605 CCGTTCATGAATATTAATTAATGAGCTTAATGAGCTTAATGAGAGAAATCCAC 10664
| | | | |
QY 3170 CCAAAACACCTTTAAAG 3186
| | | | |
Db 10665 CCAAAACACCTTTAAAG 10681
| | | | |

RESULT 3

US-09-775-398-37/C
; Sequence 37, Application US/09775398
; Patent No. 6893820
; GENERAL INFORMATION:
; APPLICANT: Plasse, Christoph
; TITLE OF INVENTION: Detection of Methylated CpG Rich Sequences Diagnostic for Maligna
; FILE REFERENCE: 22727/04075
; CURRENT APPLICATION NUMBER: US/09/775,398
; CURRENT FILING DATE: 2001-01-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens 3.B.55
US-09-775-398-37

Query Match 16.1%; Score 514; DB 3; Length 514;
Best Local Similarity 100.0%; Pred. No. 6.5e-101;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GGTGTGACATGAGACGCGCTGACCCCTGCGCCAAAGCATGCGCGGCTCGAGATCGTGC 1563
| | | | |
Db 514 GGTGTGACATGAGACGCGCTGACCCCTGCGCCAAAGCATGCGCGGCTCGAGATCGTGC 455
| | | | |
QY 1564 CCGGGGCGCGCTTACAGCATGCGCGGCTCTTTCTGTGAGAGCAATTCGTGACCAATGAC 1623
| | | | |
Db 454 CCGGGGCGCGCTTACAGCATGCGCGGCTCTTTCTGTGAGAGCAATTCGTGACCAATGAC 395
| | | | |
QY 1624 CCAAGGCGCGCTTACAGCATGCGCGGCGCGGCGCGGCGCGGCTGACAGCCGACGCG 1683
| | | | |
Db 394 CCAAGGCGCGCTTACAGCATGCGCGGCGCGGCGCGGCGCGGCTGACAGCCGACGCG 335
| | | | |
QY 1684 TGCCTGACACCAAGCGGCTGTGCGCGAGAGCGAGCGAGGACCGCGGCGCGGCTGCGC 1743
| | | | |
Db 334 TGCCTGACACCAAGCGGCTGTGCGCGAGAGCGAGGACCGCGGCGCGGCTGCGC 275
| | | | |
QY 1744 CGCAACGCTGTTTGAAGCGCGGCGCAACACCGGCTGAGCTTTCGCGGCTGCGGCTTATG 1803
| | | | |

Db 274 CGCAACGCTGTTTGAAGCGCGGCGCAACACCGGCTGAGCTTTCGCGGCTGAGCTATG 215
| | | | |
QY 1804 ACACGGCCAGGAACTTTGCGCGGCGCAACCGGCGCAAGCGCTCTCTAGCGCGCGCGGCGG 1863
| | | | |
Db 214 ACACGGCCAGGAACTTTGCGCGGCGCAACCGGCGCAAGCGCTCTCTAGCGCGCGGCGG 155
| | | | |
QY 1864 TGAAGGCGCTGCGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGG 1923
| | | | |
Db 154 TGAAGGCGCTGCGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGG 95
| | | | |
QY 1924 ACCGCTGAGGCTGAGGCGCGCGCGAGTCCCGCGAGTACTGCGGACCAAGTGGGCTCGG 1983
| | | | |
Db 94 ACCGCTGAGGCTGAGGCGCGCGCGAGTCCCGCGAGTACTGCGGACCAAGTGGGCTCGG 35
| | | | |
QY 1984 TGCCTGCTGCTGCGCGCAACAGCGCGCGGCGCGC 2017
| | | | |
Db 34 TGCCTGCTGCTGCGCGCAACAGCGCGCGGCGCGC 1

RESULT 4

US-09-513-999C-885
; Sequence 885, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, J.B.
; APPLICANT: Ductert, A.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent .pm
; SEQ ID NO 885
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 117..386
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4
; OTHER INFORMATION: w=a or t
US-09-513-999C-885

Query Match 12.1%; Score 384.6; DB 3; Length 388;
Best Local Similarity 99.7%; Pred. No. 3.9e-73;
Matches 384; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 190 TTTGCTGTTGAAGTCTTCTGTCTAGTAGAGGAGTCTGTGATTTCTAGTTTATGATA 249
| | | | |
Db 4 TTTGCTGTTGAAGTCTTCTGTCTAGTAGAGGAGTCTGTGATTTCTAGTTTATGATA 63
| | | | |
QY 250 AATAGACTTTAAACCAAGGAGCGGAGGCGAGTCTTACGTTCTAGAGCTATGACG 309
| | | | |
Db 64 AATAGACTTTAAACCAAGGAGCGGAGGCGAGTCTTACGTTCTAGAGCTATGACG 123
| | | | |
QY 310 TGAAGCACTGCTTCTCTCTATACAGCTCTCAAGAAATTTCTCAATGTGACAGCA 369
| | | | |
Db 124 TGAAGCACTGCTTCTCTCTATACAGCTCTCAAGAAATTTCTCAATGTGACAGCA 183
| | | | |
QY 370 GCTACCAACTTACAGGCGAGTCCAGCTTGTCTTGAAGATCATCCATTAATCTGACCA 429
| | | | |
Db 184 GCTACCAACTTACAGGCGAGTCCAGCTTGTCTTGAAGATCATCCATTAATCTGACCA 243
| | | | |
QY 430 CTGACAACTTGAAGAAATTTCACTTTGAAAAAATTTACAGGGGAGTACGAAATCACT 489
| | | | |
Db 244 CTGACAACTTGAAGAAATTTCACTTTGAAAAAATTTACAGGGGAGTACGAAATCACT 303
| | | | |

QY	QY	Db	QY	Db
490	CAGATACAGCAATTTTCTGACTCAAGGACCTCACCGGGGAGCTCAAGAAATTAAC	549	550	TCTCTCTCTGTTGGACGGGGTCTC
304	CAGATACAGCAATTTTCTGACTCAAGGACCTCACCGGGGAGCTCAAGAAATTAAC	363	364	TCTCTCTCTGTTGGACGGGGTCTC
				388

RESULT 5

```

US-09-188-811-1
: Sequence 1, Application US/09188811
: Patent No. 6037148
: GENERAL INFORMATION:
: APPLICANT: Khodadoust, Mehran
: TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: NNI-046CP
: CURRENT APPLICATION NUMBER: US/09/188,811
: CURRENT FILING DATE: 1998-11-09
: EARLIER APPLICATION NUMBER: 09/163,116
: EARLIER FILING DATE: 1998-09-29
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
:
: LENGTH: 1704
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1) ..(1701)
US-09-188-811-1

```

Query Match	11.8%	Score 374.6;	DB 3;	Length 1704;
Best Local Similarity	68.8%	Pred. No. 9.1e-71;		
Matches 515; Conservative	0;	Mismatches 234;	Indels 0;	Gaps 0;

QY	912	GCACAGGTGTACCTCTGTGCACAGGCCCTTTGGCTGTGAATTTACCCGGACCAAAACGGAG	971
Db	355	GCCCCAGCTTACCTGTGCACCGGGCTCTGTGGCTCAAAATTCACCGCCACCAACATGAG	444
QY	972	ATGATCATCACCAACAGGGAAGCGCATGTTCTCTTTTAAAGTTTAACTTTCCTGT	1031
Db	445	ATGATCATTTACCAAAACAGGCGACGGCGCATGTTCTCTTCTTGAAGTTCAATAAACGA	504
QY	1032	CTGCATCCACGGCTCATTTACAAATTTTTGTGATGTGATTTTGGCGATCCCAATAC	1091
Db	505	CTCAATCCACACGCCACATCACTAAGTGTTCGTGAAGGTGTGTGTGGCGAACCCCAACAC	564
QY	1092	TGGAGGTTTCAAGGAGGCGAAATGGGTTCCCTTCCGGCAAAAGCGGACCAACATGTGCAAGA	1151
Db	565	TGGCGCTTTCAGGGGGGCGAAATGGGTGACCTGTGGCAAAACCGCAATATACATGACGGC	624
QY	1152	AATCGGCTTATATGCATCCGAGTTCCGCCAACACTGSGGCTCACTGTAGTGGCCAGA	1211
Db	625	AACAAATATATGTTCACCCAGAGTCTCTAATATCTGTGTCCCACTGGATGACAGAG	684
QY	1212	ATCTCTTTGGAAATTTAAACTTTCGGAACAAAGGAGCTTCAATATACATGGGAG	1271
Db	685	ATTTCAATCGGAAATTTAAATCTCACCAATTAACAAAGCGCAATATACACACACCGAG	744
QY	1272	ATGTGTGTTTACAGTCTCTTGCAACAAGTACAGCCCGCTGSCATGTGTGTAAGTAAAC	1331
Db	745	ATGATAGTCTTCAATCTTACAAATACCAAACTCCGACCTGCATATATGTTGAAGTTACA	804
QY	1332	GAGGACGGACGGAGGACATGACGACGCCCGGCGGCTGACAGCTGTACTTTCCTGAG	1391
Db	805	GAGATGGGGGTGAGGACTTGAATGAGCCCTCAAAAGACCCAGACTTTTACTTCTCAGAA	864
QY	1392	ACTCAGTTATCGCGCTCACCGCTTACAGAAACAGGATATTTACAACTGAATAATGAT	1451
Db	865	AGGCATTTATTCAGATGACTGCTTACCAAAACAGGATATTTATCTAATTAAGATGAT	924

QY	1452	CACAACCCCTTTGGCAAAAGGATTTGGGATPAATTATGACACGATCTACACCGGCTGTGAC	1511
Db	925	CATAACCCCTTTGGCAAAAGGCTTGAGAGACACATGATTCATGTAACCGCTTCAGAA	984
QY	1512	ATGAGACGCGCTGACCCCGCTCGCCACGACTCGCGCGCTCGCAGATCGTCCCGGGGCC	1571
Db	985	AATACACAGTTTAATCTCATCTCCCACGGATTTCTCTGATGCCATCGATTTGTCCCTGSA	1044
QY	1572	CGTACGCATATGCGCGGCTCTTTCTCTGACGAGCACGTTGCTGAGCAATCTACCGAAGGCC	1631
Db	1045	GGTGGTATCGGCGCTTCAATCTCTTTCCCGGAGGCCCTTTGTCAACACTTAACTCAAGCC	1104
QY	1632	CGCTTCCACCCGGGCGCGGCGCGCGGCC	1660
Db	1105	CGCTATTATTAATGGGAGAGAACCGTGC	1133

RESULT 6

```

US-09-189-760-1
Sequence 1, Application US/09189760
Patent No. 6031078
GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehran
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: NNI-046CP2
CURRENT APPLICATION NUMBER: US/09/189,760
CURRENT FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 09/163,116
EARLIER FILING DATE: 1998-09-29
EARLIER APPLICATION NUMBER: 60/089,467
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: (PENDING)
EARLIER FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1

```

Query Match	11.8%	Score 374.6;	DB 3;	Length 2494;
Best Local Similarity	68.8%;	Pred. No. 1e-70;		
Matches 515; Conservative	0;	Mismatches 234;	Indels 0;	Gaps 0;

Qy	912	GCACAGTGTA	CTGTGTGACAGAGCCCTTGTGCTGAATTTCA	CCGGCACCAACGGAG	971
Db	398	GCCACGCTAC	CTGTGCACCGGGCTGTGGCTCAATTC	CAACCGCACCAAACTGAG	457
Qy	972	ATGATCATC	CAACCAAGGAAAGCGCATGTTTCTTTTAA	AAATTTAACATTTCGT	1031
Db	458	ATGATCATC	CAACCAAGGAGCGGACATGTTTCTTTCTTGAC	CTTCACATTAACGA	517
Qy	1032	CTGCATCC	ACGGGCTATTACATATTTTGTGATGATTTTGGCGS	ATCCCAATC	1091
Db	518	CTGCATCC	ACGCTACCAATGTGTTCTGTAAGATGTGTGCGG	ACCCCAACAC	577
Qy	1092	TGGAGATT	CAAGAGGCAAAATGGGTTCTTTCGGCAAAAGCGGA	CACCAATGTGCCAAGA	1151
Db	578	TGGCGTT	CAAGGGGGCAAAATGGGTGACCTGTGGCAAAAGCGGA	CAATTAATGCAAGGC	637
Qy	1152	AATCGGGT	CTAATATGCATCCGGATTCCCCCAAC	CTGGGGCTCATCTGATGCGCGAAGA	1211
Db	638	AACAAAT	ATATGTTCA	CCAGAGCTCTTAATCTGGTCCCATCTGATGAGACAGAG	697
Qy	1212	ATCTCTTT	TGGAAATTA	AACTTACGAACCAAGAGAGTTCAATTAACAATGGGAG	1271
Db	698	ATTTCA	TTCGGGAAATTA	AACTCAACATACCAAGGCGCAATTAACAACACACCTG	757

Oy	1212	ATGAGTGGTTTAAACAGTCCCTTTCGACAAGTACAGCCCCCGCTGCATGTGTGGAAAGTGAAC	1331
Db	758	ATGATAGCTTTACATCAATCCTTACACAAATACAAACCCCGACATGATTTGTTGAAGTTACA	817
Oy	1332	GAGGACGGCACGGAGAGACATACAGCCAGCCCGCGCGCTGCAGACGGTCACTTCCCTGAG	1391
Db	818	GAGATGGCGGTGGAGGACTTGAATAGAGCCCTCMAAGAACCCAGACTTTTACCTTCTCAGAA	877
Oy	1392	ACTCAGTTCAATCGCCGTACCGGCTACCGAACAAGATATTACAACTGAAATATGAT	1451
Db	878	AACGCAATTCATATGACAGTACGCTTACAAACACCGATATTACTCAACTAAAGATTGAT	937
Oy	1452	CACAAACCTTTTGCAAAGAAGATTTCGGATATTATGACAGATCTACACCGGCTGTGAC	1511
Db	938	CATAACCCCTTTGCMAAAGGCTTCAGAGACAACTATGATTTCCATGTATACACCGCTTCAGAA	997
Oy	1512	ATGACCGCCCTGACCCCCCTGGGCCCAACGACTCGCCGCGCCTGCGAGATGTCGCCCGGGGCC	1571
Db	998	AATGACAGGTTAACTTCATCTCCACCGAATTCCTTAGATTCCTCATGATTTGTCCCTGGA	1057
Oy	1572	CGCTACGGCATGGCGCGGCTCTTTCCTGACAGACCGATTCGTAGACAACTACGCCAAGGCC	1631
Db	1058	GATGGGTACGGGGTTCAATCCTCTTTCGCGAGGCCCTTTGTCAACACTTTACCTCAGACC	1117
Oy	1632	CGCTTCCACCCGGGCGCGGGCGCGGGGCC	1660
Db	1118	CGCTATTATTAATGGCGAGAGAACCGTGGC	1146

RESULT 7

```

US-09-514-422-1
: Sequence 1, Application US/09514422
: Patent No. 6291193
: GENERAL INFORMATION:
: APPLICANT: Rhodadoust, Mehran
: TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: MNI-046CP2
: CURRENT APPLICATION NUMBER: US/09/514,422
: CURRENT FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: US/09/189,760
: PRIOR FILING DATE: 1998-11-10
: PRIOR APPLICATION NUMBER: 09/163,116
: PRIOR FILING DATE: 1998-09-29
: PRIOR APPLICATION NUMBER: 60/089,467
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: (PENDING)
: PRIOR FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 2494
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (164)..(1714)
US-09-514-422-1

```

Query Match	11.8%;	Score 374.6;	DB 3;	Length 2494;
Best Local Similarity	68.8%;	Pred. No. 1e-70;		
Matches 515;	Conservative 0;	Mismatches 234;	Indels 0;	Gaps 0;

QY 912 CCACAGGCTGTACTGTGTGACAAAGGCCCTTTGGCTGAAATTTACCGGCACCAAAAGGAG 971

Db 398 GCCACGCTACCTGTGTCAACCGGCTCTGTGGCTGCAATTTCCACCGGCACCAAACTGAG 457

QY 972 ATGATCATCACCAAAACAGGAGGCGCATGTTCTCTTTTAAAGTTTAACTTTCTGT 1031

Db 458 ATGATCATTTAGGAACAGGCGCATGTTCTTTCTTGAGCTTCAACATATAAGGA 517

QY 1032 CTCGATCCACGGGCTCAATTACATATTTTGTGATGTGATTTGGCGGATCCCAATCAC 1091

Db	518	CTCATATCCCATCTGCCCATCTACACATATGTTCTGTAGAGTGTGTCTGGCGAATCCCAACAC	577
Qy	1092	TGAGAGTTTCAAGAGAGGCAATGGTTCTCTCGGCAAAAGCGACACCAATGTGCAAGA	1151
Db	578	TGGCCCTTCAAGGGGGGGCAATAGGGTGACCTGTGGCAAAAGCGACCAATGACATGACGGGC	637
Qy	1152	AATGGGCTATATGATATCCGAGTTCCGCCAACATCGGGGGCTCATCTGATATGGGCCAAGA	1211
Db	638	AACAAAATGATATGTTCAACCCAGATCTCTTAATATCTGGTTCCACATGATGACACAGAG	697
Qy	1212	ATCTCTTTGGAAAAATTAATAACTTACGAAACAACAAAGAGCTTCAATATTAACAATGGGCA	1271
Db	698	ATTTCAATCGGGAATTTAAATCACTACCAATTAACAAGGCGCAATTAACAACAACACCCAG	757
Qy	1272	ATGTGTTTACAGTCTCTTGACAAGTACAGCCCGCCTGATGTGGTAGTGAAC	1331
Db	758	ATGATAGTCTTCAATATCCTTACACAATAACCAACCCCGCATGTATTTGTGAAGTTACA	817
Qy	1332	GAGGACGGCAGGAGAGACCTAGCCAGCCCGGGCGGTGCAGACGTTCACTTTCCCTGAG	1391
Db	818	GAGGATGGGTGAGAGACTTGAAATGAGCCTTCAAGAACCCAGACTTTACCTTCTCAGAA	877
Qy	1392	ACTCAGTTTCATGCGCGGTCAACCGCTTACACAGAACAGCGATATTACAACACTGAAATATAGT	1451
Db	878	AACGATATTCATTTGACGTACTGTCTTACCAAAACACGATATTTACTCACTAAAGATTGAT	937
Qy	1452	CACAACTCTTTTGCAAAAGAGATTTGGGATTAATATGACACGATCTTACACCGGCTGTAC	1511
Db	938	CATAACCCCTTTGCAAAAGGCTTCAGAGACAATATGATTCATATGACACCGGCTTCAGAA	997
Qy	1512	ATGACCGCGCTGACCCCGTCGCGCCAAACGACTGCGCGCGCTCGAGATCGTGCCCGGGGCC	1571
Db	998	AATGACAGGTTAATCTTCATCTTCCACAGATTTCTCTAATATCCCATAGATTTGCTCCTGGA	1057
Qy	1572	CGCTACGCCATGCGCGGCTCTTTCTGTGACAGACAGATTCGTGACAACTACGCCAAGGCC	1631
Db	1058	GGTGGTACGGGTTCAATCTCTTTCCCGAGGCCCTTTGTCAACACTTACTCAAGCC	1117
Qy	1632	CGCTTCCACCGGGGGCGGGCGCGGGGCC	1660
Db	1118	CGCTATTTAATGCGAGAGAACGTTGCC	1146

RESULT 8

```

US-09-949-016-3434
/ Sequence 3434, Application US/09949016
/ Patent No. 6012339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3434
/
/ LENGTH: 2575
/
/ TYPE: DNA
/
/ ORGANISM: Human
/ US-09-949-016-3434

```

Query Match	9.9%;	Score 316.2;	DB 3;	Length 2575;
Best Local Similarity	61.2%;	Pred. No. 3.8e-58;		
Matches 510; Conservative	0;	Mismatches 323;	Indels 0;	Gaps 0

696 CCGGCGACCCCGCCTTCTCCATCGGCGAGCCCTAGCCGTTACATGGCCACCAACCCGGTC 755

Db	404	CCCGCCCGCGAGCCGCTTCTCTTGAGACTTAACGCTTACCGCGCGACCCAGGCGGC	463
Qy	756	ATCAACAAAGAGCCTTACAAAGACCTCTCTGTCAACTCTCGCGAGAGATACCCACG	815
Db	464	GAGTTTCCCGGGGCGGGGAGAGTCTTCCCGCGCCCGGAGACGCGAGGGCTACAGCG	523
Qy	816	GCCGCTACCCCTTACCCACAGAGTACGGCCCTCTTACCAAGAGCTCCGTTTACAG	875
Db	524	GCGGAGGAGCTACGCGCGCCCGGACCCCGCGCCCGGCTCTTACCGGCGCGGTAGAGC	583
Qy	876	TTCTCTCCACCCAGCCGCGGGCTGGTGCCTCCGCAAGCAGAGTGTACTCTGTCAACAG	935
Db	584	TACGCGCTAACCCGCGGAGCTGAGAGTGTGGGAAAACTGAGAGTGGCGTTCAACACAC	643
Qy	936	CCCCCTTGGCTGAATTTTCAACCGGACCAAAAGAGATGATCATCAACAAAGGAAAG	995
Db	644	CTGTTGTGAGTCCCAAGTTTATCAGCACAGACAGAGATCATCATCAACAGAGGACG	703
Qy	996	CGCATGTTTCTTTTAAAGTTTAACTTTCTGCTCTGATCCCAAGGCTATTTACAT	1055
Db	704	CGGATGTTTCCCAATTCCTGTCAATTTACTGTGGCGGGCTGGAACCCACAGCACTTACAG	763
Qy	1056	ATTTTGTGAGATGATATTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAAAATG	1115
Db	764	AGTTTGTGAGAGCTGTGCTTGTGTGACACAGCACCACTGCGGTTACAGAGCGCAAGTGG	823
Qy	1116	GTTCTCTGCGCAAAAGCGGACCAATGTGCAAGAAATCGGAGTCTATATGCAATCGAGT	1175
Db	824	GTCGACTGTGAAAGGCGAGGCGAGCAGCAGCAGGAAACCGCTGTACGTCCACCCGAGC	883
Qy	1176	TCGCCCAACACTGGGGGCTCACTGTGATGCGCCCAAGAAATCTTTTGGAAAAATTAACTT	1235
Db	884	TCCCCCAACACAGAGCGCACTGAGTGCCTCAGAAAGTTTCAATTTGGGAAACTTAAAGCTC	943
Qy	1236	AGCAACAAAGAGAGCTTCAATTAACAATGGGCGAGTGTGTTTACAGTCTTGAC	1295
Db	944	ACAAACAACAAAGGGGCGTCAACAATGTGATCCAGATATTTGTCTTCCAGTCCCTCCAT	1003
Qy	1296	AAGTACCAAGCCCCGCTGATGTGTGGAAGTGAACGAGAGCGGACGAGGACACTAGC	1355
Db	1004	AAGTACCAAGCCCCGCTGATATCTGTGTGAGTGAACGAGAGGACGAGGAGCTGCG	1063
Qy	1356	CAGCCCGCGCGGCTGCAACGTTCACTTTCCTGAGACTCACTTCATGCGCCGTACCGCC	1415
Db	1064	AACGCTTCCAAACGCAATATCTTACTTTCCTCAAGAAACCAAGTTCATTCCTGACTGCGC	1123
Qy	1416	TACCAAGAACGAGATTTTACACACTGGAATATGATCAACAACCTTTTGCAAAGAGATT	1475
Db	1124	TACCAAGAAAGCGAGATTACTCAGCTGAATATGATTAATTAACCCCTTGTCCAAAGGATTC	1183
Qy	1476	GGGAGTAATTTATGACAGATCTACACCGGCTGTGACATGAGACCGGCTGACCC	1528
Db	1184	CGGAGAACTTTGATGTCACTATACATCTGTGTAACCAAGCATCCCTCTCC	1236
<p>RESULT 9 US-949-016-976 Sequence 976, Application US/09949016 Patent No. 6812339 GENERAL INFORMATION: APPLICANT: VENTER, J. Craig et al. TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF FILE REFERENCE: C1001307 CURRENT APPLICATION NUMBER: US/09/949, 016 CURRENT FILING DATE: 2000-04-14 PRIOR APPLICATION NUMBER: 60/241, 755 PRIOR FILING DATE: 2000-10-20 PRIOR APPLICATION NUMBER: 60/237, 768 PRIOR FILING DATE: 2000-10-03 PRIOR APPLICATION NUMBER: 60/231, 498 PRIOR FILING DATE: 2000-09-08</p>			

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO: 976
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-976

```

Query Match	9.9%	Score 316.2;	DB 3;	Length 2589;
Best Local Similarity	61.2%;	Pred. No. 3.8e-58;		
Matches 510;	Conservative 0;	Mismatches 323;	Indels 0;	Gaps 0

OY	696	CCGGGAGACCCCGGCTTCTCCATCGGAGGCGCTAGCGGCTAATAGGCGCCACACCGCGATC	755
Db	404	CCCGCCCGCGGAGCCCGCTTCTTGGAGCTTACGCTTACCCCGCGGACCCCGAGCGGCC	463
OY	756	ATCAACCAACGAGGAGCTTCAACAGCGCTCTGTCCAACTCTCGCGGAGGATACCCACG	815
Db	464	GAGTTCCTCCCGCGCGGGGAGAGTCTTTCGCGCGCCGCGAGCGCCGAGGAGCTACAGCG	523
OY	816	GCCGGCTAACCCCTTACCCACAGAGTACGGCCACTCTCTCAACAGGAGCTCGTTCACG	875
Db	524	GCGCAGGAGCTTACGCGCGCCCGGAGCCCGCGCCGCGGCTTACCCGCGGCGCGTGAGAGC	583
OY	876	TTTCTCTTCAACGACCGGAGCTGTGCGCCCGGCAAGACACAGTGTACTTGTGCACAG	935
Db	584	TACGGCTACCCCGCGGAGCTGAGAGTGTGGGGGAAATGAGAGGTTCGGGCTCAACACAC	643
OY	936	CCCTTTGGCTGAATAATTACCGGACCAACGAGAGTATCATCAACCAACAGGAGG	995
Db	644	CTGTGTGGTCCAAAGTTTAATCAACACAGACAGAGATGATCATCAACCAAGGACGG	703
OY	996	CGCATGTTCTCTTTTAAATTAACTTTCTGTCTCATCCCACGGCTCATTTACAAT	1055
Db	704	CGAGTGTCCCATTCCTGTCAATTAACTGTGCGCGGCTGAGGCCACACGCGCTTACG	763
OY	1056	ATTTTGTGTGATGTGATTTTGGCGGATCCCAATCATGAGAGGTTTCAAGAGCAATATG	1115
Db	764	ATGTTTGTGAGCGTGTCTTGTGTGACGACACATCGCGGTATACAGAGCGGCAATGTG	823
OY	1116	GTTCCTTTCGGCAAGCGGACCAATGTGCAAGAAATCGGATCTATATGATCCGGAT	1175
Db	824	GTGCAGTGTGAAAGCCGAGGGACGATGCAAGAAACCGCTGTACGCCACCGGAC	883
OY	1176	TCCCCCAACTGTGGGCTCACTGATGCGCCAAAGAAATCTTTTGGAAATTTAAACTT	1235
Db	884	TCCCCCAACAGAGGCGCACTGTGATGCGCAGAGAAAGTTTCAATTGGGAAACTAAAGCTC	943
OY	1236	ACGAACAAACAAGAGCTTCAATTAACAATGGGAGATGATGGTTTTACAGTCTTGAC	1295
Db	944	ACAAACAAACAAGGGGCGCTCAACAACTGTGACCCAGATGATTTGTCTCAATCTCTCAT	1002
OY	1296	AAGTACACGAGCCCGCTCATGTGTGTGAATGTGAACGAGACCGGACCGAGACACTAGC	1355
Db	1004	AAGTACACGAGCCCGCTCATGTGTGTGAATGTGTGAACGAGAGCCAGAGGACGCTGCG	1062
OY	1356	CAGCCCGGCGGCGTGCAGACGTTCACTTTCCCTGAGACTCACTTTCATGTCCGCTACCGCC	1415
Db	1064	AACGCTTCCAACACGCAATCTTTTCACTTTCCAAAGAAACCGATTCATTTCCGTGACTGCC	1122
OY	1416	TACCGAACAACGAGATATTACAACAACGTGAATGATGTCACACCCCTTTGGCAAAAGATTT	1475
Db	1124	TACCGAATGCGGAGATTACTAGTGAATATGATTAATACCCCTTTGCCAAAGATTC	1182
OY	1476	CGGATTAATTATGACAGCATCTACACCGGCTGTGACATGAGACCGCTGACCC	1528
Db	1184	CGGAGAACTTTGAGTCCATGTACACATCTGTGTGACACAGAGATCCCTGCC	1236


```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 6783961
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent .pm
; SEQ ID NO 22243
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-22243
```

```
Query Match          7.4% Score 235; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 4.7e-41;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      2923 TTTTATTCGTCAAGGACAAACAGTTCACTTAACTTTTTCCTTTCT
Db      1   TTTTATTCGTCAAGGACAAACAGTTCACTTAACTTTTTCCTTTCT 60
Oy      2993 TTGCTTTCTTCTCTCTCTATCTTCTCTCTCTTTAAATTTTCTGTGGA 3042
Db      61   TTGCTTTCTTCTCTCTCTATCTTCTCTCTCTTTTAAATTTTCTGTGGA 120
Oy      3043 TAAATATTCAGAGGCTCTAGAAACATGAAATCTCAGTAGAGTGGTTTCCACTTCT 3102
Db      121 TAAATATTCAGAGGCTCTAGAAACATGAAATCTCAGTAGAGTGGTTTCCACTTCT 180
Oy      3103 CCTCAATCCGTGATGAAATATTAATTAATGCTTATGACACAAATAGCTA 3157
Db      181 CCTCAATCCGTGATGAAATATTAATTAATGCTTATGACACAAATAGCTA 235
```

```
RESULT 11
US-09-949-016-656
; Sequence 656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 656
; LENGTH: 1798
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-656
```

```
Query Match          4.7% Score 150.6; DB 3; Length 1798;
Best Local Similarity 58.4%; Pred. No. 1.3e-22;
Matches 330; Conservative 0; Mismatches 214; Indels 21; Gaps 3;
```

```
Oy      919 TGTACTGTGCAACGAGCCCTTTGGCTGAATTTTCCACCGGACCAACGAGATGATCA 978
Db      302 TGAGCTGTGAGAACCGGAGCTATGAGAGAGATTGAGCTGTGTGGAAACAGAAATGATCA 361
```

```
Oy      979 TCACCAAAAGGAGGAGCGCATGTTCTTTTAAAGTTTAAACATTTCTGCTCGATC 1038
Db      362 TCACCAAAAGCTGGAGAGCGCATGTTCTTCTGCTCCGAGTGTACGTCACTGGCCTGGACC 421
Oy      1039 CCAGGCTCATTAATATTTTGTGATGATTTTGGCGGATCCCAATCATCTGGAGT 1098
Db      422 CCGAGGCCCGCTACTTGTGTTCTTCTGTGATGATTTCCGGTGGATGGGGCTGCTACCGCT 481
Oy      1099 TTCAAGAGCAAAATGGGTTCTTGGCGCAAGGCGACCAATGTCAGAAATGGG 1158
Db      482 GGCAGGGCGGCGCTGGAGCCAGCGCAAGGAGCCCGCTGCTG--ACGTC 538
Oy      1159 TCTATATGATCCGATTTCCCAACACAGGCGCTCATGAGTGGCCCAAGAAATCTT 1218
Db      539 TCTATATGATCCGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
Oy      1219 TTGAAATTTAAATTTAATGCAACCAAGAGCTTCAATATACATGGGAGATGGTCG 1278
Db      599 TCCATCGTGTCAAGCTCAACCAAGAGAGCTGAGACCCCAAGGCA-----CCTGA 649
Oy      1279 TTTTACAGTCTTGGACAGTACAGCCCGGCTGCAATGTTGGAAGTGAAGAGAGC 1338
Db      650 TCTGCACTCCATGACAGTACCAACCCGCTATACCTAGTTGCGGACGCCAGCTCT 709
Oy      1339 GCACGAGACACTAGCCAGCCCGCGCTGACAGCTTCACTTCCCTGAGACTGAGT 1398
Db      710 GCACGAG-----CACTGGGGGGCAATGGCTCTCTTCCGCTCCGAGACCAAT 760
Oy      1399 TCATGCGGTCACCGCTTACCAACAGGATTTTACACATGAAATATGATACAAAC 1458
Db      761 TCATGCGGTCACCGCTTACCAACAGGATTTTACACATGAAATATGACCAATC 820
Oy      1459 CTTTGAAGAGATTTGCGGATTA 1483
Db      821 CTTTGAAGAGCTTCCGGGAGAA 845
```

```
RESULT 12
US-09-949-016-2908
; Sequence 2908, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2908
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2908
```

```
Query Match          4.7% Score 150.6; DB 3; Length 1799;
Best Local Similarity 58.4%; Pred. No. 1.3e-22;
Matches 330; Conservative 0; Mismatches 214; Indels 21; Gaps 3;
```

```
Oy      919 TGTACTGTGCAACGAGCCCTTTGGCTGAATTTTCCACCGGACCAACGAGATGATCA 978
Db      302 TGAGCTGTGAGAACCGGAGCTATGAGAGAGATTGAGCTGTGTGGAAACAGAAATGATCA 361
Oy      979 TCACCAAAAGGAGGCGCATGTTCTTTTAAATTTTAACTTTTCTGCTCGATC 1038
Db      362 TCACCAAAAGCTGGAGAGCGCATGTTCTTCTGCTCCGAGTGTACGTCACTGGCCTGGACC 421
```


1039 CCACGGCTCATTAACATATTTTGTGATGTGATTTGGCGGATCCCATATCATCTGAGAGT 1098
 422 CCGAGGCCCGCTACTTGTCTTCTTGATGTGATTCGGGTGATGAGGCTCGCTACCCCT 481
 1099 TTCAAGAGGCAAAATGGGTTCTTGGCGCAAAAGCGGACCAATGTGCAAGGAATTCGGG 1158
 482 GGCAGAGGCGGCGCTGGAGGCCAGCGCAAGAGGCGCCGCTGCTG---ACCGTG 538
 1159 TCTATATGATCCGAGATTCCTCCCAACACTGGGGCTCATCTGATGTGCGCAAGAAATCTCTT 1218
 539 TCTATATCACTCCGACATCTCTGCGCACTGTGTCAATTTGATGTGGAGGAGCCTGTGTCTT 598
 1219 TTGAAAATTTAAATTAAGTAAAGCAAAAGAGCTTCAATTAACAATGAGGAGATGTGG 1278
 599 TCCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
 1279 TTTTACATGCTTGTGCAAGTACCAAGCCCGCTGATGTGTGGAAGTGAAGGAGAGCG 1338
 650 TCTGCACTCCATGACAGATGACCAAGCCCGCATACCTAGTTGGGAGGCCAGCTT 709
 1339 GCACGAGGACACTAGCCAGCCGCGGTGCAAGCTTCACTTCCCTGAGACTCAAT 1398
 710 GAGCCAG-----CACTGGGGGGGATGAGCTCTTCCGCTTCCCGAGACCAAT 760
 1399 TCATGCGCTGCAAGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1458
 761 TCATGCGCTGCAAGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 820
 1459 CTTTGGCAAAAGATTTCCGGATTA 1483
 821 CTTTGGCAAAAGATTTCCGGAGAA 845

RESULT 13
 US-09-949-016-4821
 ; Sequence 4821, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 4821
 ; LENGTH: 2433
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-4821

Query Match 4.6%; Score 146; DB 3; Length 2433;
 Best Local Similarity 57.2%; Pred. No. 1.4e-21;
 Matches 335; Conservative 0; Mismatches 230; Indels 21; Gaps 3;

899 GGTGCGCGGCAAGCAAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTTCACCG 958
 555 GGTGAGAGAGACCCCAAGGTGACCTGTGAGGCTTAAGAACTTTGGATCACTTTTCAAA 614
 959 GCACCAAGAGAGATGATCATCAACCAAGAGAGGAGCGATGTTTCTTTTAAAGTT 1018
 615 GCGGGGCAAGAGATGATCATTAACCAAGTGGGAAAGCGAATGTTTCTCCATTTAAAGT 674
 1019 TAAATTTCTGTGTGATCCAGGCTCATTAATTTTGTGATGTGATTTTGGC 1078

675 GAGATGTTCTGGCGCTGATTAATAAAGCCAAATACATTTATGTGATGATTAATGACTGC 734
 1079 GATATCCCAATCATCTGAGGTTTCAAGAGGCAATGGGTTCTTCCGCAAAAGCGACAC 1138
 735 TGAATGATGCTGATTAATTAATTTCAATTTCTGGATGATGTGGCTGTAAAGCCGACCC 794
 1139 CAATGTCGAAGGAATGCGGTCTATATGATCCGATTCCTCCCAACTGGGGCTCACTG 1198
 795 GGAATGCG---CAAGAGGATGATCATTCACCCGAGAGGCCCTCTACTGGGGAACAGTG 851
 1199 GATGCGCAAGAAATCTCTTTTGAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1258
 852 GATGTCGAAGTCTGATTTTCCAAACTGAACTCAACCAACAACTTTAG----- 904
 1259 TAAATGAGGAGATGTGTGTTTACAGTCTTTCGCAAGATGACAGCCCGCTGATGT 1318
 905 --ACAAATGATGATTTACTATATTAATGATGATGATGATGATGATGATGATGATGAT 962
 1319 GGTGAAATGAAAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1378
 963 TGTAAAGCCCAATGATCATTTGAAAT-----CCCTATATGATCATTTCCGACATTA 1013
 1379 CACTTTCCTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 1014 CTGTTTCCCGAAACTGATTAATTCATGCTGTGATGATGATGATGATGATGATGATGATGAT 1073
 1439 ACTGAATATGATCAACCTTTTGAATAAGATTTGGGAGTAA 1484
 1074 GTTAAATATGACCAACACCTTTTGAATAAGATTTCCGGAGACT 1119

RESULT 14
 US-09-949-016-4822
 ; Sequence 4822, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 4822
 ; LENGTH: 2433
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-4822

Query Match 4.6%; Score 146; DB 3; Length 2433;
 Best Local Similarity 57.2%; Pred. No. 1.4e-21;
 Matches 335; Conservative 0; Mismatches 230; Indels 21; Gaps 3;

899 GGTGCGCGGCAAGCAAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTTCACCG 958
 555 GGTGAGAGAGACCCCAAGGTGACCTGTGAGGCTTAAGAACTTTGGATCACTTTTCAAA 614
 959 GCACCAAGAGAGATGATCATCAACCAAGAGAGGAGCGATGTTTCTTTTAAAGTT 1018
 615 GCGGGGCAAGAGATGATCATTAACCAAGTGGGAAAGCGAATGTTTCTCCATTTAAAGT 674
 1019 TAAATTTCTGTGTGATCCAGGCTCATTAATTTTGTGATGTGATTTTGGC 1078
 675 GATATGTTCTGGCTGATTAATAAAGCCAAATTAATTTATGATGATTAATTAATTAATTAAT 734
 1079 GATATGTTCTGGCTGATTAATAAAGCCAAATTAATTTATGATGATTAATTAATTAATTAAT 1138

```

Db      735 TGATACCTGCTTATTAATTTACAAATTCGCGTGTGATGTGCGCTGTGTAAGCCGACCC 794
Qy      1139 CAATGTGCAAGAAATCGGGCTCTATATGATCCGATTCGCCAACACTGGGGCTCACTG 1198
Db      795 CGAAATGCG---CAAGAGAGATGTACATTCACCCGACAGCCCGCTACTGGGGGAACAGTG 851
Qy      1199 GATGGCGCAAGAAATCTCTTTTGGAAATTTAACTTTGCAACACACAAAGAGCTTCAAA 1258
Db      852 GATGTCCAAAGTCGTCACTTTCCACAAACTGAACTCACCAACAACTTTTCAAG----- 904
Qy      1259 TAACAATGGCGCAGATGTGTGTTTTCAGTCTTTGCAAGTACAGCCCGCTGTCATGT 1318
Db      905 --ACAAACATGATTTACTATATTTGAATCCATGCACAAAATACAGCCCGCTTCCACAT 962
Qy      1319 GGTGAAGTGAACGAGAGACGGCAGGAGACACTAGCCAGCCCGCGCTGACAGCTT 1378
Db      963 TGTAGAGCCATATGACATCTTGAACCT-----CCCTTATAGTACATTTCCGACATA 1013
Qy      1379 CACTTCCCTGAGACTGAGTTCATGCGCGTCACGGCTACAGAGACAGGATATTACACA 1438
Db      1014 CTGTGTTCCCGAACTGATTTATGCTGTGTGACTGATACAGAAATGATTAAGATTAACCA 1073
Qy      1439 ACTGAAATATAGATCAAACTCTTTTGCAAAAGATTTGGGATTAAT 1484
Db      1074 GTTAAATATAGCAACAACCTTTTGCAAAAGTTTCCGGGACACT 1119

```

```

RESULT 15
US-10-028-272-1
; Sequence 1, Application US/10028272
; Patent No. 6635481
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: TBX3 GENE AND METHODS OF USE
; FILE REFERENCE: 16U 104 R1
; CURRENT APPLICATION NUMBER: US/10/028,272
; CURRENT FILING DATE: 2001-12-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 3113
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (524)..(2695)
; OTHER INFORMATION:
US-10-028-272-1

```

```

Query Match      4.6%; Score 146; DB 3; Length 3113;
Best Local Similarity 57.2%; Pred. No. 1.6e-21;
Matches 335; Conservative 0; Mismatches 230; Indels 21; Gaps 3;

```

```

Qy      899 GGTGCCCGGCAAGCAGACGCTGTAACCTGTGCAACAGGCCCTTTGGCTGAAATTTACCG 958
Db      817 GGTGAGAGACGACCCCAAGGTGACCTGAGGCTAAAGAACTTTGGATCACTTTACAA 876
Qy      959 GCACCAAAAGAGATGATCATCAACCAACAGGAAAGCGCATGTTCTTTTAAAGTTT 1018
Db      877 GCGGGGACACGAGATGTGTCATTAACCAAGTCGGGAAGCGAATGTTTCTCCATTTAAAGT 936
Qy      1019 TAACTTTTGTGTCTGATCCACGCGCTCATTAACAATATTTTGTGATGTATTTGGC 1078
Db      937 GAGATGTTCTGGCTGTGATTAAGAAAGCCAAATACATTTTATGATGACATTATAGCTGC 996
Qy      1079 GGATCCCAATCACTGGAGGTTTCAAGGAGGCAAAATGGGTTCTTGGCGCAAGCGACAC 1138
Db      997 TGATGACTGTCTGTATTAATTTTCACAATTTCTGGTGTGATGTGGCTGTAGGCCGACCC 1056
Qy      1139 CAATGTGCAAGGAAATCGGGCTCTATATGATCCGATTCGCCCAACACTGGGGCTCACTG 1198
Db      1057 CGAAATGCG---CAAGAGAGATGTATCATTCACCCGAGACAGCCCGCTACTGGGGAAACAGTG 1113

```

```

Qy      1199 GATGCGCAAGAAATCTCTTTTGGAAATTTAACTTACGAAACAACAAGAGCTTCAAA 1258
Db      1114 GATGTCCAAAGTGTCTCACTTTCCACAACTGAAATCTACCAACAACATTTTCAAG----- 1166
Qy      1259 TAACAATGGCGCAGATGTGTGTTTTCAGTCTTTGCAAGTACAGCCCGCTGTCATGT 1318
Db      1167 --ACAAACATGATTTTACTATATTTGAATCCATGCACAAAATACAGCCCGCTTCCACAT 1224
Qy      1319 GGTGAAGTGAACGAGAGACGGCAGGAGACACTAGCCAGCCCGCGCTGACAGCTT 1378
Db      1225 TGTAGAGCCATATGACATCTTGAACCT-----CCCTTATAGTACATTTCCGACATA 1275
Qy      1379 CACTTCCCTGAGACTGAGTTCATGCGCGTCAACGCTACAGAAACAGGATATTACACA 1438
Db      1276 CTGTGTTCCCGAACTGATTTATGCTGTGTGACTGATACAGAAATGATTAAGATTAACCA 1335
Qy      1439 ACTGAAATATAGATCAAACTCTTTTGCAAAAGATTTGGGATTAAT 1484
Db      1336 GTTAAATATAGCAACAACCTTTTGCAAAAGTTTCCGGGACACT 1381

```

```

Search completed: December 20, 2005, 16:05:30
Job time : 548 secs

```


Db 61 CAGGAGCCGTTTCATCGCGTCGACAGACGACAGATCAAAAGTGACCTTCTGATTCG 120
Qy 121 TGCATAGTGTCAATTGGCCCATCTCTTCTCCGAGGAAAAAAGTAAATCAAACTT 180
Db 121 TGCATAGTGTCAATTGGCCCATCTCTTCTCCGAGGAAAAAAGTAAATCAAACTT 180
Qy 181 TGAGAGAGATTTGCTGCTGTTGAAGTCTTCTGCTAGTGAAGGGGGTCTGTGATTTCTAG 240
Db 181 TGAGAGAGATTTGCTGCTGTTGAAGTCTTCTGCTAGTGAAGGGGGTCTGTGATTTCTAG 240
Qy 241 TTTATGATTAATAGACTTTTAAAAACAAGGAGCGGAGGCGAGTTCAGGTTCTAGAG 300
Db 241 TTTATGATTAATAGACTTTTAAAAACAAGGAGCGGAGGCGAGTTCAGGTTCTAGAG 300
Qy 301 CTATGAGCTGGAGACTGCGCTTCTCTCTATGATGCTCTCAAGAAATTTCTCAATG 360
Db 301 CTATGAGCTGGAGACTGCGCTTCTCTCTATGATGCTCTCAAGAAATTTCTCAATG 360
Qy 361 TGAGAGAGAGCTAACCCACATTCAGGCGGATCCGAGCTTGTCTTGACATCCCATTA 420
Db 361 TGAGAGAGAGCTAACCCACATTCAGGCGGATCCGAGCTTGTCTTGACATCCCATTA 420
Qy 421 TCTCGACCACTGACAACTTGAGAGAAAGTTCACTTTGAAAAAATTAACAAGGGAGATGA 480
Db 421 TCTCGACCACTGACAACTTGAGAGAAAGTTCACTTTGAAAAAATTAACAAGGGAGATGA 480
Qy 481 CGAATCACTGATGACATACAGACAAATTTTCTGACTCCAAGAGCTCACAGGGGAGCTCCAGA 540
Db 481 CGAATCACTGATGACATACAGACAAATTTTCTGACTCCAAGAGCTCACAGGGGAGCTCCAGA 540
Qy 541 GAAAGTAACTCTCTCTGCTTGAGAGGGGGTCTGAGAGCTTGATGACAGTTTCAGTGGT 600
Db 541 GAAAGTAACTCTCTCTGCTTGAGAGGGGGTCTGAGAGCTTGATGACAGTTTCAGTGGT 600
Qy 601 CTGCTGAGAGTGTCT 660
Db 601 CTGCTGAGAGTGTCT 660
Qy 661 CCAAGTGCATGTTTCCCGTACCCCGGCGAGACAGGACCGGCGACCCCGCTTCTCTCATG 720
Db 661 CCAAGTGCATGTTTCCCGTACCCCGGCGAGACAGGACCGGCGACCCCGCTTCTCTCATG 720
Qy 721 GGAAGCCCTTAGCGGCTCATATGAGCCACACCGGGTATCAACAAGGAGCTTACACAGCC 780
Db 721 GGAAGCCCTTAGCGGCTCATATGAGCCACACCGGGTATCAACAAGGAGCTTACACAGCC 780
Qy 781 TCTCTGCAACTCTCTGCGCGAGGGATACCCACCGCGGGCTTACCCTTACCCACAGCAGT 840
Db 781 TCTCTGCAACTCTCTGCGCGAGGGATACCCACCGCGGGCTTACCCTTACCCACAGCAGT 840
Qy 841 ACGGCGCACTCTTACCAAGAGAGCTCGGTTTACCAAGTTCTCTCTCACCCAGCCGGGGCTGG 900
Db 841 ACGGCGCACTCTTACCAAGAGAGCTCGGTTTACCAAGTTCTCTCTCACCCAGCCGGGGCTGG 900
Qy 901 TGCAGCGGCAAGACAGGAGTACCTGTGACAGGAGCCCTTGGCTGAAATTTCAACCGGC 960
Db 901 TGCAGCGGCAAGACAGGAGTACCTGTGACAGGAGCCCTTGGCTGAAATTTCAACCGGC 960
Qy 961 ACCAAACGAGATGATCATACCAACAAGAGAGGCGATGTTTCTTTTAAAGTTTAA 1020
Db 961 ACCAAACGAGATGATCATACCAACAAGAGAGGCGATGTTTCTTTTAAAGTTTAA 1020
Qy 1021 ACAATTTCTGCTCGATCCACGGCTCATTTACAAATTTTGTGAGTGAATTTGGCGG 1080
Db 1021 ACAATTTCTGCTCGATCCACGGCTCATTTACAAATTTTGTGAGTGAATTTGGCGG 1080
Qy 1081 ATCCCAATGACTGAGAGTTTCAAGAGGAGAAATGGTTCTTGGCGCAAAAGGGACACA 1140
Db 1081 ATCCCAATGACTGAGAGTTTCAAGAGGAGAAATGGTTCTTGGCGCAAAAGGGACACA 1140
Qy 1141 ATGTGCAAGAAATCGGGTCTATATGATTCGGAATTCGCCCAACATGGGGGCTCACTGGA 1200
Db 1141 ATGTGCAAGAAATCGGGTCTATATGATTCGGAATTCGCCCAACATGGGGGCTCACTGGA 1200

Qy 1201 TGGCCCAAGAAATCTCTTTGAAAAATTAAAACTTACGAACAACAAGAGCTTCAATA 1260
Db 1201 TGGCCCAAGAAATCTCTTTGAAAAATTAAAACTTACGAACAACAAGAGCTTCAATA 1260
Qy 1261 ACAATGGGACAGATGGTGGTTTACAGTCTTGCACAGTACACAGCCCGCTGCATGG 1320
Db 1261 ACAATGGGACAGATGGTGGTTTACAGTCTTGCACAGTACACAGCCCGCTGCATGG 1320
Qy 1321 TGAAGTGAACGAGAGCGGACGAGAGACATAGCCAGCCCGCGCTGCAGCTTCA 1380
Db 1321 TGAAGTGAACGAGAGCGGACGAGAGACATAGCCAGCCCGCGCTGCAGCTTCA 1380
Qy 1381 CTTTCCCGAGAGCTCAGTTATGCGCGTCAACCGGCTTCCAGAACAGGATTTACACAC 1440
Db 1381 CTTTCCCGAGAGCTCAGTTATGCGCGTCAACCGGCTTCCAGAACAGGATTTACACAC 1440
Qy 1441 TGAATAATGATCACAACTTTTGCAAAAGATTTGCGGATTAATTAAGACAGATCTACA 1500
Db 1441 TGAATAATGATCACAACTTTTGCAAAAGATTTGCGGATTAATTAAGACAGATCTACA 1500
Qy 1501 CCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAACAGACTGCGCGCTCGCATGCG 1560
Db 1501 CCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAACAGACTGCGCGCTCGCATGCG 1560
Qy 1561 TGGCCCGGGCGCGCTTACGCAAGGCGGCTCTTCTGAGAGACAGTGGTGAGAACT 1620
Db 1561 TGGCCCGGGCGCGCTTACGCAAGGCGGCTCTTCTGAGAGACAGTGGTGAGAACT 1620
Qy 1621 ACGCAAGGCGCGCTTCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGA 1680
Db 1621 ACGCAAGGCGCGCTTCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGA 1680
Qy 1681 GCGTGCAGCACCAAGAGGCTGTGTGTGCGCGAGAGGCGGAGGAGACCCGGGCGGCGCT 1740
Db 1681 GCGTGCAGCACCAAGAGGCTGTGTGTGCGCGAGAGGCGGAGGAGACCCGGGCGGCGCT 1740
Qy 1741 GCGCGCAACGCTGGTTTGTGAGCGCGGCGCAACCGGCTGGAATTTGCGGCGCTCGGCT 1800
Db 1741 GCGCGCAACGCTGGTTTGTGAGCGCGGCGCAACCGGCTGGAATTTGCGGCGCTCGGCT 1800
Qy 1801 ATGACACGGCCACGAGACTTCCGCGGCAACCGGCGCACGCTGCTTACCGGCGGCGG 1860
Db 1801 ATGACACGGCCACGAGACTTCCGCGGCAACCGGCGCACGCTGCTTACCGGCGGCGG 1860
Qy 1861 GCGTGAAGGCGCTGCGCTGAGGCTGAGGCTGACCTGCGCGCGCTGCGCTACTACG 1920
Db 1861 GCGTGAAGGCGCTGCGCTGAGGCTGAGGCTGACCTGCGCGCGCTGCGCTACTACG 1920
Qy 1921 CCGAGCCGCTGGGGCGGCGGCGGCGAGTCCCGCGAGTACTGCGGCAACAGTCGAGGCT 1980
Db 1921 CCGAGCCGCTGGGGCGGCGGCGGCGAGTCCCGCGAGTACTGCGGCAACAGTCGAGGCT 1980
Qy 1981 CCGTGTGCTGCTGCTGCGCAACAGCGCGCGGCGCGCGCATGAGGCTCGGCGCAATC 2040
Db 1981 CCGTGTGCTGCTGCTGCGCAACAGCGCGCGGCGCGCGCATGAGGCTCGGCGCAATC 2040
Qy 2041 CTTACTTGGCGAGAGGCGGAGGCTGTGCTGCGCGAGCGCTGCGCTGCGCGCGGCG 2100
Db 2041 CTTACTTGGCGAGAGGCGGAGGCTGTGCTGCGCGAGCGCTGCGCTGCGCGCGGCG 2100
Qy 2101 CCGCGGAGGAGCGGCAAGCGGCAAGGACCTGTCCGATTCAGCTGAGTGAAGCGGCTCT 2160
Db 2101 CCGCGGAGGAGCGGCAAGCGGCAAGGACCTGTCCGATTCAGCTGAGTGAAGCGGCTCT 2160
Qy 2161 CGATCAAGTGCATGACTCCAGCACTCGGAGATTTAGAGCAAGCGGAGCGGAGG 2220
Db 2161 CGATCAAGTGCATGACTCCAGCACTCGGAGATTTAGAGCAAGCGGAGCGGAGG 2220
Qy 2221 TCTGCGCGGCGCAACGCGCGTGTCCGAGATTTGCTCCGCTCAAGAGCGAGGCTGCG 2280
Db 2221 TCTGCGCGGCGCAACGCGCGTGTCCGAGATTTGCTCCGCTCAAGAGCGAGGCTGCG 2280

2281 CCCAGCGGAGCTGCGAGAAAGAACTGCGCCAGAGCAATTAGCGGCTACTAGGCTTCTACT 2340
 2281 CCCAGCGGAGCTGCGAGAAAGAACTGCGCCAGAGCAATTAGCGGCTACTAGGCTTCTACT 2340
 2341 GCGACAGTGGGCG 2400
 2341 GCGACAGTGGGCG 2400
 2401 CCTCAGCTCTTCCCGAGCTCGCGCTCGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
 2401 CCTCAGCTCTTCCCGAGCTCGCGCTCGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
 2461 ATTGACCTCGATGCGCGCTCTGCGAGCAATTAAGTGCAGTCTCCGAGCGTAAATTTTAC 2520
 2461 ATTGACCTCGATGCGCGCTCTGCGAGCAATTAAGTGCAGTCTCCGAGCGTAAATTTTAC 2520
 2521 CTTTTTGGACAGCGCTCTGCAATTAAGTGCAGTCTCCGAGCGTAAATTTTAC 2580
 2521 CTTTTTGGACAGCGCTCTGCAATTAAGTGCAGTCTCCGAGCGTAAATTTTAC 2580
 2581 TTTGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
 2581 TTTGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
 2641 GTCTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
 2641 GTCTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
 2701 TGGGAGCTCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 2760
 2701 TGGGAGCTCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 2760
 2761 CATAGATGTTGACTCTAGAACCTGAGCCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2820
 2761 CATAGATGTTGACTCTAGAACCTGAGCCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2820
 2821 GATG 2880
 2821 GATG 2880
 2881 ATG 2940
 2881 ATG 2940
 2941 ACAAACAAGTCTATGCTTAACTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 3000
 2941 ACAAACAAGTCTATGCTTAACTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 3000
 3001 TCTCATACTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3060
 3001 TCTCATACTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3060
 3061 TAGAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3120
 3061 TAGAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3120
 3121 AATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 3121 AATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 3181 TAAAGG 3186
 3181 TAAAGG 3186
 3181 TAAAGG 3186

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 FILE REFERENCE: ABOMICA-X-2
 CURRENT FILING DATE: 2001-12-20
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 20697
 LENGTH: 864
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC009487.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
 OTHER INFORMATION: NT HIT: 91142971, EVALUATE 0.00e+00
 OTHER INFORMATION: SWISSPROT HIT: Q16550, EVALUATE 0.00e+00
 OTHER INFORMATION: EST_HUMAN HIT: AI201654.1, EVALUATE 0.00e+00
 US-10-029-386-20697
 Query Match 27.1%; Score 864; DB 6; Length 864;
 Best Local Similarity 100.0%; Pred. No. 1,4e-230; Indels 0; Gaps 0;
 Matches 864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1493 GATCTACACCGGCTGTGACATGAGCCGCTGACCCCTCGCCCAAGACTGCGCGCTC 1552
 1 GATCTACACCGGCTGTGACATGAGCCGCTGACCCCTCGCCCAAGACTGCGCGCTC 60
 1553 GCAGATCGTCCCG 1612
 61 GCAGATCGTCCCG 120
 1613 GAGCAACTACGCGCAAGGCG 1672
 121 GAGCAACTACGCGCAAGGCG 180
 1673 GAGCCGCGAGCGTGCAGCAACCAACGAGCGCTGTCGCGCGAGCAAGCGCGAGCGCGCGCG 1732
 181 GAGCCGCGAGCGTGCAGCAACCAACGAGCGCTGTCGCGCGAGCAAGCGCGAGCGCGCGCG 240
 1733 GCGCGCTCGCGCGCAACGCTGTTGTGACGCGCGCGCAACCGCGCTGAGACTTGGCGCG 1792
 241 GCGCGCTCGCGCGCAACGCTGTTGTGACGCGCGCGCAACCGCGCTGAGACTTGGCGCGCG 300
 1793 CTCGCGCTATGACAGCGCGCAACGCTGTTGTGACGCGCGCGCAACCGCGCTGAGACTTGGCG 1852
 301 CTCGCGCTATGACAGCGCGCAACGCTGTTGTGACGCGCGCGCAACCGCGCTGAGACTTGGCG 360
 1853 GCGCGCGCGCGTGAAGCGCTGCG 1912
 361 GCGCGCGCGCGTGAAGCGCTGCG 420
 1913 CTACTACGCGCAACCGTGGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
 421 CTACTACGCGCAACCGTGGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 1973 GTGCGGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2032
 481 GTGCGGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
 2033 CGCAATCCCTACCTGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2092
 541 CGCAATCCCTACCTGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 2093 GCCCGGCGCGCGCGAGAGCGCGCAAGCGCGCAAGCACTGTCGATTCAGCTGATGAGAGC 2152
 601 GCCCGGCGCGCGCGAGAGCGCGCAAGCGCGCAAGCACTGTCGATTCAGCTGATGAGAGC 660
 2153 GCCCTCTCGATCAAGTCAATGAGCTCCAGCGAGCGAGGATTTAGCAAGCGCGCAAGCG 2212

RESULT 2
 US-10-029-386-20697
 Sequence 20697, Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.

QY	DB	Query Match	Best Local Similarity	Matches	Score	DB	Length	Indels	Gaps
QY	661	US-10-363-345A-9001/c	80.5%	768	604	8	1026	6	5
QY	2213	Sequence 9001, Application US/10363345A							
DB	721	Publication No. US20040234960A1							
QY	2273	GENERAL INFORMATION:							
DB	781	APPLICANT: Alexander Olek							
QY	2333	APPLICANT: Kurt Berlin							
DB	841	TITLE OF INVENTION: Method for determining the degree of methylation of defined							
		FILE REFERENCE: E01/1227							
		CURRENT APPLICATION NUMBER: US/10/363,345A							
		CURRENT FILING DATE: 2003-03-03							
		NUMBER OF SEQ ID NOS: 40712							
		SEQ ID NO 9001							
		LENGTH: 1026							
		TYPE: DNA							
		ORGANISM: Artificial Sequence							
		FEATURE:							
		OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)							
		OTHER INFORMATION: Cpg-island No: 9001							
		US-10-363-345A-9001							
QY	1490	US-10-363-345A-9001/c	80.5%	768	604	8	1026	6	5
DB	949	Sequence 9001, Application US/10363345A							
QY	1550	Publication No. US20040234960A1							
DB	889	GENERAL INFORMATION:							
QY	1610	APPLICANT: Alexander Olek							
DB	829	APPLICANT: Kurt Berlin							
QY	1670	TITLE OF INVENTION: Method for determining the degree of methylation of defined							
DB	771	FILE REFERENCE: E01/1227							
QY	1730	CURRENT APPLICATION NUMBER: US/10/363,345A							
DB	711	CURRENT FILING DATE: 2003-03-03							
QY	1790	NUMBER OF SEQ ID NOS: 40712							
DB	652	SEQ ID NO 9001							
QY	1850	LENGTH: 1026							
DB	592	TYPE: DNA							
QY	1910	ORGANISM: Artificial Sequence							
DB	1910	FEATURE:							
		OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)							
		OTHER INFORMATION: Cpg-island No: 9001							
		US-10-363-345A-9001							

Dd		533 CGACTCTACAGCCGACCCTGTGAATTAABAAOCCCCGCAATTCCC GGCAATACTAAGAACGC	474
Oy		1970 CAAGTCGGGCGTTCGTGTGCCCTGTGTGSCCCAAGCQCQCGCGCCGCGCATGCG	2029
Dd		473 CAAATTCGAATCTGTAATCT - CCTTAATAACCAAACCGCGCGACCGCGCGGATAAAC	415
Oy		2030 GCGCGCCAAATCCCTTTCTTGCGCGAGAAGAGCCGAGGSGCTTGCGCGCGGAGCGCTCGCGCT	2089
Dd		414 GAGCGCCCAATCTCTTACTTAACAAGAAAAACGAAAACCTTAACGCGGAACGCTCGCGCT	355
Oy		2090 GCCCGCCGCGCGCCGCGAGAGCACGCGCAAGCCCAAGAGCCTGTTCGATTCCAGCTGTGATCGA	2149
Dd		354 ACCGCCC - GCGCGCGCGAAAAGCCGAAACCCGAAAACCTATTCGATFTCCAACTTAATGSA	296
Oy		2150 GAOCGCTCTTCGTGATCAAGTCATCGATCTTCAGCGACTTGGGGAATTACAGACAGGCCAA	2209
Dd		295 AACGCCCTCTTCGATCAATCATCATCGATCTCCAAACGACTGAAAAATTATGTAACMAACCA	236
Oy		2210 GCGAGGCGGATCTTCGCGCGCGCGAAGCCCGTGTGCCAGATTGTCCCGGCTCAAGAG	2269
Dd		235 ACGAAAAAGAAATCTGCGCGACCGACGACGCGGTATCCGAAAATTGTCCCGCTCAAAA	176
Oy		2270 CGAGGTGCTGGCGCCGAGCGGACCTGCGAGAAAGAACTGCGCAAGACATTAGCGGCTACTTA	2329
Dd		175 CGAAATACTTAACCCCAAGAACTACGAAAAAAAACTTAGCGCAAAAATTAAAGACTACTTA	116
Oy		2330 TGGCTTCTACTTCGCACAGCTNAGGCGCCCTGTGCGCGCGCGCGCGCGCGCGCGAGCC	2389
Dd		115 TAAC TTCTACTGCGACAACTAAACCGCCCTTACCGCGCGACCGCGCGCGCGCGGAAAC	56
Oy		2390 CCCAGCCAGCCCTCTCAAGCTCTTCCCAAGCTCGCGCTCCGCACTCCTCTT	2443
Dd		55 CCACAACAAACCCCTCAACAATCTTCCCAACTCGCGCTCCCACTCCTCTT	2
<hr/>			
RESULT 4			
US-10-363-345A-9002			
; Sequence 9002, Application US/10363345A			
; Publication No. US20040234960A1			
; GENERAL INFORMATION:			
; APPLICANT: Alexander Olek			
; APPLICANT: Christian Piependroek			
; TITLE OF INVENTION: Method for determining the degree of methylation of defined			
; FILE REFERENCE: EOI/1227			
; CURRENT APPLICATION NUMBER: US/10/363,345A			
; NUMBER OF SEQ ID NOS: 40712			
; SEQ ID NO 9002			
LENGTH: 1026			
TYPE: DNA			
ORGANISM: Artificial Sequence			
FEATURES:			
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)			
OTHER INFORMATION: CpG-island No: 9002			
US-10-363-345A-9002			
<hr/>			
Query Match 19.0% ; Score 604; DB 8; Length 1026;			
Best Local Similarity 80.5%; Pred. No. 9.9e-158;			
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;			
<hr/>			
Oy		1490 CACGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTTGCGCCCAAGACTGCGCGG	1549
Dd		78 CAAATCTACACCGACTTAATTAATTAACCGCTTAAACCCCTTGCGCCCAAGACTGCGCGG	137
Oy		1550 CTGCGAGATCGTCCCGGGGCGCGCTACATGCGCGGCTCTTTCTGAGAGCACTT	1609
Dd		138 CTCGCAATCTGTAACCGAAACCGCTACGCACTTAACCGACTTTTTCTTCAAAACCAATT	197
Oy		1610 CGTGAGCAACTAGCGCAAGCGCGGCTTCACCGCGGCGGCGGCGCGCGCGCGG	1669
Dd		198 GTTAACAACATAAGCCAAACCGGCTTCCACCGCAAGG - GACGGAACCGGAACCGAA	255

1670 TACGAGCGGACGCGTGGCGGACACCAAGGCGTGTGCGCGCGACAGCGCGGAGACCC 1729
 256 TACGAAACGCGACGTAACCGCACCAACGAACTACTATGCGCGCAACAAACCGCAACCC 315
 1730 GGGCGCGCGCTTGGCGGACGCTGAGTTGTGAGCGCGCGGACCAACCGGCTGAGCTGGC 1789
 316 GAAACGG-CCTGCGCGCAACGCTAATTATTAAGCGCGACCAACCGGCTAATCTTGGC 374
 1790 GCGCTGCGGCTATGACACGCGGACGAGCTTGGCGGCAACGCGGCGACGCTGCTCTTA 1849
 375 GACCTGACCTATTAACACGACCACTTCCGGAACAAAGGACGACGACTCTCTTA 434
 1850 GCGCGCGCGGCGGTGAAGCGCTGCGGTGAGGCTGAGCTGAGCTGGCGCGCGCT 1909
 435 CGGACGACGAAAGTAAACGCT-ACGCTAACAACTAACAACTAACCGCGCGCT 493
 1910 GCGCTATGACGCGACCGTGGGCTGGGCGCGCGGACGCTGCGGACGCTGAGCGAC 1969
 494 CGACTATGACGCGACCGTGGGCTGGGCGCGGACGCTGCGGACGCTGAGCGAC 553
 1970 CAAGTGGGCTGCGTGGCGGCTGCGGCAACGAGCGCGCGCGCGGACGCTGAGCG 2029
 554 CAATGAACTGATCT-CCCTACTAACCGCAACGCGCGGACGCGCGCGGCTGAGCG 612
 2030 GCGCGCAATCCCTACTGCGGAGGAGCGGCGGCGGCTGCGCGCGGAGCGCTGCGCGCT 2089
 613 CGACGCAATCCCTACTGCGGAGGAGCGGCGGCGGCTGCGCGCGGAGCGCTGCGCGCT 672
 2090 GCGCGCGGCGCGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2149
 673 ACCGCGG-CGCGCGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 731
 2150 GAGCGCGCTGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2209
 732 AAGCGCGCTGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 791
 792 ACGAAACGAAATTCGCGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 851
 2210 GCGGAGCGGAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2269
 2270 CGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2329
 852 CGAAATATCTAACCGCAACGAAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 911
 2330 TGGCTTCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2389
 912 TAACTTCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 971
 2390 CCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2443
 972 CCGAACCAACCGCTGACGAGCTTCCCGCACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1025

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; OTHER INFORMATION: Cpg-island No: 9001
 US-10-363-483A-9001
 Query Match 19.0%; Score 604; DB 9; Length 1026;
 Best Local Similarity 80.5%; Pred. No. 9.9e-158;
 Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

1490 CACGATCTACACGCGCTGTGACATGAGCGCGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGG 1549
 949 CAAATCTACACGCGCTATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 890
 1550 CTGCGAGATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1609
 889 CTGCGAGATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 830
 1610 GGTGAGCAATACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1669
 829 CGTAAACAACTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 772
 1670 TACGAGCGGACGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1729
 771 TACGAGCGGACGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 712
 1730 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1789
 711 GAAACGG-CCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 653
 1790 GCGCTGCGGCTATGACACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1849
 652 GACCTGACCTATTAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 593
 1850 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1909
 592 CGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 534
 1910 GCGCTATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1969
 533 CGACTATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
 473 CAATGAACTGATCT-CCCTACTAACCGCAACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415
 2030 GCGCGCAATCCCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2089
 414 CGACGCAATCCCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355
 2090 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2149
 354 ACCGCGG-CGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 296
 2150 GAGCGCGCTGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2209
 295 AAGCGCGCTGCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 236
 2210 GCGGAGCGGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2269
 235 ACGAAACGAAATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 176
 2270 CGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2329
 175 CGAAATATCTAACCGCAACGAAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 116
 2330 TGGCTTCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2389
 115 TAACTTCTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 56
 2390 CCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2443
 55 CCGAACCAACCGCTGACGAGCTTCCCGCACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2

RESULT 6
US-10-363-483A-9002
; Sequence 9002, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9002
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 9002
US-10-363-483A-9002

Query Match 19.0%; Score 604; DB 9; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9,9e-158;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

QY 1490 CACGATCTACACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 1549
DB 78 CAAATCTACACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 137
QY 1550 CTGCGAGATGATGACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 1609
DB 138 CTGCGAGATGATGACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 197
QY 1610 CGTGAACAATACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 1669
DB 198 CGTGAACAATACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 255
QY 1670 TACGAGACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 1729
DB 256 TACGAGACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCCGCCG 315
QY 1730 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 1789
DB 316 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 374
QY 1790 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 1849
DB 376 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 434
QY 1850 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 1909
DB 435 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 493
QY 1910 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 1969
DB 494 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 553
QY 1970 CAAAGTGGGCTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2029
DB 554 CAAAGTGGGCTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 612
QY 2030 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2089
DB 613 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 672
QY 2090 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2149
DB 673 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 731
QY 2150 GAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2209

DB 732 AACGCGCTCTGATCAATTCATGACCTCCAGACTCGAAAAATTAGAACAAACCA 791
QY 2210 GCGAGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2269
DB 792 GCGAGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 851
QY 2270 GCGAGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2329
DB 852 GCGAGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 911
QY 2330 TGGCTTCTACTGCGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2389
DB 912 TGGCTTCTACTGCGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 971
QY 2390 CCGAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 2443
DB 972 CCGAGCGGCGGTGTGACATGAGACCGGTGTGACATGAGACCGGTGTGACATGAGACCGGT 1025

RESULT 7
US-10-029-386-6986
; Sequence 6986, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6986
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009487.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: H10108.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q16650, EVALUATE 5.00e-15
; OTHER INFORMATION: NT HIT: g111429971, EVALUATE 0.00e+00
US-10-029-386-6986

Query Match 16.2%; Score 515; DB 6; Length 515;
Best Local Similarity 100.0%; Pred. No. 5,6e-133; Indels 0; Gaps 0;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2236 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 2295
DB 1 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 60
QY 2296 AGAAGAACTGGCGCAAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 2355
DB 61 AGAAGAACTGGCGCAAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 120
QY 2356 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 2415
DB 121 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 180
QY 2416 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 2475
DB 181 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 240
QY 2476 GCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGTGTGAGAGATTCGCGCGGT 2535

Db 241 GCCGTGACAGGAATTAAGTGAAGTCTCCGAGGCGAATTTAACTTTTTCACACA 300
QY 2536 GTCTTGAATTAAGTCAACCGAAGCTTCACTTTGCTGTAACCTTTTGTCTTACT 2595
Db 301 GTCTTGAATTAAGTCAACCGAAGCTTCACTTTGCTGTAACCTTTTGTCTTACT 360
QY 2596 ACTCTTCTTGTGAGATTATCTCTCAATTTCCCTCCCTGTCTTCTTACT 2655
Db 361 ACTCTTCTTGTGAGATTATCTCTCAATTTCCCTCCCTGTCTTCTTACT 420
QY 2656 CCTACTTCTTCTTGTGTAATAAAGCTTCACTTTGAGAGACCTGGGAGTCTGTCA 2715
Db 421 CCTACTTCTTCTTGTGTAATAAAGCTTCACTTTGAGAGACCTGGGAGTCTGTCA 480
QY 2716 GGCAGCAGGATTCGAGCCCGCAAGTCTGGGCT 2750
Db 481 GGCAGCAGGATTCGAGCCCGCAAGTCTGGGCT 515

RESULT 8
US-10-842-072-37/c
Sequence 37, Application US/10842072
Publication No. US20050158731A1
GENERAL INFORMATION:
APPLICANT: PLASS, CHRISTOPH
TITLE OF INVENTION: DETECTION OF METHYLATED CPG RICH SEQUENCES DIAGNOSTIC FOR
FILE REFERENCE: 22727-04242
CURRENT APPLICATION NUMBER: US/10/842,072
CURRENT FILING DATE: 2004-05-10
PRIOR APPLICATION NUMBER: 09/775,398
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.2
SEQ ID NO 37
LENGTH: 514
TYPE: DNA
ORGANISM: Homo sapiens
US-10-842-072-37

Query Match 16.1%; Score 514; DB 9; Length 514;
Best Local Similarity 100.0%; Pred.No.11e-132; Indels 0; Gaps 0;

Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GCGTGAATGAGACCGCTGACCCCTCCGCAAGACTCGCGCTCGAGATGTC 1563
Db 514 GCGTGAATGAGACCGCTGACCCCTCCGCAAGACTCGCGCTCGAGATGTC 455
QY 1564 CCGGGCCCGCTAAGCAGTGGCGGCTCTTCTGAGAGACAGTTCGTGAGCACTACG 1623
Db 454 CCGGGCCCGCTAAGCAGTGGCGGCTCTTCTGAGAGACAGTTCGTGAGCACTACG 395
QY 1624 CCAAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683
Db 394 CCAAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
QY 1684 TGCAGACACCAAGGAGTGTGTCGCGCGAGAGCGGAGAGCGGAGCGCGCGCGCG 1743
Db 334 TGCAGACACCAAGGAGTGTGTCGCGCGAGAGCGGAGAGCGGAGCGCGCGCGCG 275
QY 1744 CCGAAGCTGTTGTGACCGCGCGCAACACCGGCTGACCTTTCGCGGCTCGGCTTATG 1803
Db 274 CCGAAGCTGTTGTGACCGCGCGCAACACCGGCTGACCTTTCGCGGCTCGGCTTATG 215
QY 1804 ACAAGCGGACGAGCTTTCGCGCGCAACCGCGCGAGCGGAGCGGAGCGGAGCGG 1863
Db 214 ACAAGCGGACGAGCTTTCGCGCGCAACCGCGCGAGCGGAGCGGAGCGGAGCGG 155
QY 1864 TGAAGCGGCTGCGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1923
Db 154 TGAAGCGGCTGCGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 95
QY 1924 ACCGCTGCGGCTGCGGCGCGCGAGTCCCGCGAGTACGCGGACCAAGTCCGCGCTCG 1983

Db 94 ACCGCTGCGGCTGCGGCGCGCGAGTCCCGCGAGTACGCGGAGCACCAAGTCCGCGCTCG 35
QY 1984 TGCCTCCCTGCTGGCCCAACAGCGCGCGCGCGCG 2017
Db 34 TGCCTCCCTGCTGGCCCAACAGCGCGCGCGCGCG 1

RESULT 9
US-10-363-345A-9003
Sequence 9003, Application US/10363345A
Publication No. US20040234960A1
GENERAL INFORMATION:
APPLICANT: Alexander Olek
APPLICANT: Christian Piepenbrock
APPLICANT: Kurt Berlin
TITLE OF INVENTION: Method for determining the degree of methylation of defined
FILE REFERENCE: E01/1227
CURRENT APPLICATION NUMBER: US/10/363,345A
CURRENT FILING DATE: 2003-03-03
NUMBER OF SEQ ID NOS: 40712
SEQ ID NO 9003
LENGTH: 1026
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-345A-9003

Query Match 14.7%; Score 469.6; DB 8; Length 1026;
Best Local Similarity 71.7%; Pred.No.4.2e-120; Indels 6; Gaps 5;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

QY 1491 ACGATCTACACCGGCTGATGATGACCGCGCTGACCCCTCGCCCAAGACTCGCGCGC 1550
Db 79 ACGATCTACACCGGCTGATGATGACCGCGCTGACCCCTCGCCCAAGACTCGCGCGC 138
QY 1551 TCGAGATGATGCG 1610
Db 139 TCGAGATGATGCG 198
QY 1611 GTGACAACTAAGCCAAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCG 1670
Db 199 GTGACAACTAAGCCAAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCG 256
QY 1671 ACGACCGGAGGCG 1730
Db 257 ACGACCGGAGGCG 316
QY 1731 GCGCGCGCTCGCGCGCAACCGCTGTTGTGACCGCGCGCGCGCGCGCGCGCGCG 1790
Db 317 GCGCGCGCTCGCGCGCAACCGCTGTTGTGACCGCGCGCGCGCGCGCGCGCGCG 375
QY 1791 GCGTGGCTTATGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1850
Db 376 GCGTGGCTTATGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 435
QY 1851 GCGGCG 1910
Db 436 GCGGCG 494
QY 1911 GCGTACTACCGCAACCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1970
Db 495 GCGTACTACCGCAACCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 554
QY 1971 AAGTGGGCTGCGTGTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2030
Db 555 AAGTGGGCTGCGTGTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 613
QY 2031 GCGGCAATCCCTACCTGGGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2090

```
Db      614 GGCCTTAATTTTATTTGGGCGAGGAGGTCGAGGGTTTGTGTCGAGCGTTTCGTGCTTG 673
Qy      2091 CCGCCCGGCGCCCGGAGGAGCCCAAGCCCAAGACCTGTCCGATTCCAGCTGATTCGAG 2150
Db      674 TCGTTC-GCGTCGTGAGAGACGTTAAGTTTGGTTCGATTGTTTAAAGTTGATCGAG 732
Qy      2151 ACGCCCTCCTGATCAAGTCCATTCAGACTCCAGCTCGGGGATTTCAGAGGAGCCAAAG 2210
Db      733 ACGTTTTTTCATTAAGTTTATCGATTTCAGATTACGATTCGGGGATTTCAGAGTAAAG 792
Qy      2211 CGAGCGCGGATCTCGCCGCGGCGAGCACGCCCGTGTCCGAGAGTTGCTCCCGCTCAAGAC 2270
Db      793 CGAGCGCGGATTTTCGTCGCGTTCGATACGTTTCGTTTCGAGAGTTTCGTTTAAAGAC 852
Qy      2271 GAGGTGCTGCGCCAGCGGAGACCTCGGAGAAAGACTCGGCCAAGACATTACGGCTACAT 2330
Db      853 GAGGTGTTGGTTTAAAGGAGATTGCGAGAAAGATTGCTTAAAGAAATTAAAGCGTTATAT 912
Qy      2331 GGCCTTCTACTCGCACAGTAGAGCGCCCTGCGCCGCGCCGCGCGCGGAGCC 2390
Db      913 GGTTTTATTCGTATTAAGTTTAAAGTTCGTTTTCGTTTCGTTTCGTTTCGCGGTTCCGATTT 972
Qy      2391 CGAGCGCGCCCTCAACAGCTCTTCCCGAGCTCGCCCTCCCACTCTCTCTTG 2444
Db      973 TTAGTTAGTTTATTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTTGG 1026

RESULT 10
US-10-363-345A-9004/C
; Sequence 9004, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/12327
; CURRENT APPLICATION NUMBER: US/10/363,345A
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9004
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-345A-9004

Query Match      14.7%; Score 469.6; DB 8; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;
```

```
Qy      1791 GCTTCGCTTATGACAGGGCCAGGACTTCGCGGAGCAAGCGGCGACGCTGCTTAC 1850
Db      651 GTTTCGGTTATGATACGGGTTACCGAATTCGCGGTTAACCGGTTACGTTGTTTTTAC 592
Qy      1851 GCGCGCGCGCGGCGTGAAGGCGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCTC 1910
Db      591 GCGCGCGCGGCGTGAAGGCGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCTC 533
Qy      1911 GCTTACACGCGGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1970
Db      532 GGTATTAACGTCGATTCGTCGCGGTTGCGGCGTTGTAAGTTTTCGTAAGTATTCGTAAT 473
Qy      1971 AAGTCGCGCTGCGTGCCTGCTGCGCCCAACAGCCCGCGCGCGCGCGCGCGCGCG 2030
Db      472 AAGTCGCGCTGCGTGCCTGCTGCGCCCAACAGCCCGCGCGCGCGCGCGCGCGCG 414
Qy      2031 GCGCCCAATCCCTACCTGCGGAGAGAGCCGAGGCGCTGCGCGCGAGCGCTGCGCTG 2090
Db      413 GCGCTTATTTTATTTTATTTGCGGAGAGAGTGAAGGTTTGGTCGTCGAGCGTTTCGCTTG 354
Qy      2091 CCGCCCGCGCGCGCGCGAGAGAGCCCAAGCCCAAGACCTGTCGATTCAGCTGATGAG 2150
Db      353 TCGTTC-GCGTCGTGAGAGAGTTAAGTTTAAAGATTGTCATTTTAAAGTTGATGAG 295
Qy      2151 ACGCCCTCCTGATCAAGTCCATGCACTCCAGCACTCGGGGATTTACAGACAGCCAAAG 2210
Db      294 ACGTTTTTTCGATTAAGTTTATCGATTTCAGATTTCAGATTTCAGATTTCAGATTTCAG 235
Qy      2211 CGAGCGCGGATCTCGCGCGCGAGACGCCCTGTGTCGAGATTGTCGCCCTCAAGAGC 2270
Db      234 CGAGCGCGGATTTGTCGCGTGCATACGTTTCGTTTCGAGAGATTGCTTTGTTTAAAGAC 175
Qy      2271 GAGGTGCTGCGCCAGCGGAGCTGCGAGAAAGACTGCGCCCAAGACATTAGGCTACTAT 2330
Db      174 GAGGTGTTGTTTAAAGCGGATTCGAGAAAGATTGCTTAAAGATTAAAGCGTTATAT 115
Qy      2331 GGCCTTCTACTCGCACAGTAGAGCGCCCTGCGCGCGCGCGCGCGCGCGCGAGCC 2390
Db      114 GGTTTTATTCGTATTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTGG 55
Qy      2391 CGAGCGCGCCCTCAACAGCTCTTCCCACTCGCGCTCCCACTCTCTCTTG 2444
Db      54 TTAGTTAGTTTATTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTTGG 1

RESULT 11
US-10-363-483A-9003
; Sequence 9003, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; NUMBER OF SEQ ID NOS: 2003-03-03
; SEQ ID NO 9003
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-483A-9003

Query Match      14.7%; Score 469.6; DB 9; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;
```

1491	ACAGTCTACACCGGCTGTGACATGACCGCGCTGACCCCTCGGCCAACGACTCGCCGCG	1550
1492	AGGATTAATATACGGGTGTGATATAGATACGTTGATTTTTCGTTAAACGATCGTCCGCT	1551
1493	TCGACAGATTCGAGCCCGGGGCGCGCTACGCGCATGCGCGGCTCTTTCCTGACAGACAGTTC	1552
1494	TCGTAAGATCGTGTTCGGGGTTCGTTAGCGTAATGCGCGGTTTTCGTTAGAGATTAGTTC	1553
1495	GTGAGCAACTACGCCAAGGCCCGCTTCCACCCTGGGCGTGGGCGGGGCGCCCGGGT	1554
1496	GTGAGTAAATACGTTAAGGTTTCGTTTAAATTCGGGCGCGG--CGCGGGTTCGGGTCGGGT	1555
1497	ACGACACCGACAGGTCGCGCACACCAACAGGGGCTGCGTGGCGGACAGAGCGGACCGG	1556
1498	ACGATCGACACGTCGTGCTATATTAACCGGTTGTTCGTGTAGTACGATTCG	1557
1499	GGCGCGCCTCGCGCGCAACGCTGCTTTGTGACCGCGCGCAACACCGGCTGACCTTCGCG	1558
1500	GGCGCG--TTTCGTGCTAACGTTGGTTGTGACGTCGTTAATAACGTTGATTCGCG	1559
1501	GCTTCGCGCTTATGACACGCGCAACGGAATTTCGGGGCAACCGGCGCACGCTGCTCTTAC	1560
1502	GTTTCGCTTATGATACGTTAACGATTTCCGCGGTAAACGCGTTATCCTGTTTTTTTAC	1561
1503	CGCGCGCGGCGCGCTGAAAGCGCTCCGCGTGCAGGCTGCAGGCTGACGCGCGCGCTC	1562
1504	CGCGCGCGGCGCGCTGAAAGCGCTCCGCGTGCAGGCTGCAGGCTGACGCGCGCGCTC	1563
1505	CGCGCGCGGCGCGCTGAAAGCGCTCCGCGTGCAGGCTGCAGGCTGACGCGCGCGCTC	1564
1506	GGCTACTACGCGACCGCTCGGGGCTGCGGCGCGCGCACTCCCGCGACGTAATGCGCGCAC	1565
1507	GGTTATTAACGTCGATTCGTCGGGTTGGGGCGCTGCTATTTTCGTAATTTGGCGATAT	1566
1508	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1567
1509	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1568
1510	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1569
1511	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1570
1512	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1571
1513	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1572
1514	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1573
1515	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1574
1516	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1575
1517	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1576
1518	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1577
1519	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1578
1520	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1579
1521	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1580
1522	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1581
1523	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1582
1524	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1583
1525	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1584
1526	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1585
1527	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1586
1528	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1587
1529	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1588
1530	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1589
1531	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1590
1532	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1591
1533	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1592
1534	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1593
1535	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1594
1536	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1595
1537	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1596
1538	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1597
1539	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1598
1540	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1599
1541	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1600
1542	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1601
1543	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1602
1544	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1603
1545	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1604
1546	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1605
1547	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1606
1548	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1607
1549	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1608
1550	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1609
1551	AACTCGGCGCTCGGTGCTGCGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGGCAATGCGC	1610
1552	AACTCGGCGCTCGGTGCTGCG	

	TITLE OF INVENTION:	Diagnosis of illnesses or predisposition to certain
	FILE REFERENCE:	82011
	CURRENT APPLICATION NUMBER:	US/10/363,483A
	CURRENT FILING DATE:	2003-03-03
	NUMBER OF SEQ ID NOS:	40712
	SEQ ID NO	9004
	LENGTH:	1026
	TYPE:	DNA
	ORGANISM:	Artificial Sequence
	FEATURE:	
	OTHER INFORMATION:	chemically treated genomic DNA (Homo sapiens)
	OTHER INFORMATION:	Cpg-island No: 9004
	US-10-363-483A-9004	
Query Match	14.7%;	Score 469.6; DB 9; Length 1026;
Best Local Similarity	71.7%;	Pred.No. 4.2e-120;
Matches	684; Conservative	0; Mismatches 264; Indels 6; Gaps 5;
QY	1491	ACGATCTACACCGGCTGTGCATGAGACC GGCTTAACCCCTTCGGCCAAAGCATGCGCGCGC 1550
DB	948	AGGATTATATACGGTTTGTGATATGAGATCGTTTGAATTTTTTGTTAACATTCGTCCGGT 889
QY	1551	TTCGAGATTCGGCCCGGGGCCCGCTCAAGCATGACGCGGACTCTTCTCTGACAGACAGATTC 1610
DB	888	TCGAGATTCGTTTCGGAGTTTCCTTACGTTATGATCGGTTTTTTTTTGTAAGATTAAGTT 829
QY	1611	GTVGACAATAACGCCAAGCCCCGCTTCACCCCGGACGGGACGGGACCCCGGGCCGGGT 1670
DB	828	GTGATATATTAAGTTAAGTTTCGTTTTTATTCGGGCGCG--GCGGGTTTCGGGTCCGGGT 771
QY	1671	ACGAGACCGACAGCTGCGCACACCAACAGGCTGCTGTGTCGCGACAGCGCCGAGGACCCG 1730
DB	770	ACGAGATGTAAGCTGTGTGATATTAAAGGGTTGTGTCGTAAGTAGAGGTCGAGGATTCG 711
QY	1731	GGGCGCGCCCTGCGCGCAACGCTGGTTTGATGACGCGGCAACAACGCGCTGGAATTGCGC 1790
DB	710	GGCGCG--TTTCGTGTAAACGTTGGTTTGTAAGTCGTTAATACGTTGATTTGCGC 652
QY	1791	GCCTCGCGCTTATGACACGCGCCACGGACTTCGCGGCAACGCGGACGCGCTGCTCTTTAC 1850
DB	651	GTTTCGATTATGATACGTTAACGGAATTGCGGGTAACGCGGTTACGTTGTTTTTTAC 592
QY	1851	GCGGCGCGCGGCGGTGAAGCGCTGCGCTGCGAGCTGACAGCTGCATGCGCGCGCGCTC 1910
DB	591	GCGGCGCGGCGGTGAAGCGCTG--CGTTGAGGTTGTAGGTTGTATTGGTCTGTCCTTC 533
QY	1911	GCGTACTACGCGGACCGGTGCGGCGCGCGGCGCGGACGTCCC CGGACGTACTGCGGCAAC 1970
DB	532	GGTATTAACGTGACATTCGTGCGGTTTGCGGCGCTTGTAAGTTTTTGTAAGTATTCGGTAAT 473
QY	1971	AAGTCGCGGCTCGAGCTGACCTGCTGTCGGCCCAACGACGCGCGCGCGCGCATGGCC 2030
DB	472	AAGTCGCGGCTCGAGT--TTTTTTGTGTGTTAATAGCGTCGCGGTGTCGCGCATGGCTC 414
QY	2031	GCGCGCAATCTCTTACTGTGCGAGAGGCGGAGGCGCTGCGCGCGACGCTCGCGCTG 2090
DB	413	GCGCTTAATTTTATTGCGGCGAGAGAGGTGAGGAGTTTGTCGTCGACGCTTCGTCGTG 354
QY	2091	CGCGCGCGCGCGCGGAGAGCCAGCCAGGACCTGTGTCGATTCAGCTGAGTCGAGTCGAG 2150
DB	353	TCGTTTC--GCGTCGTCGAGAGGATTAAAGTATTTGTCGATTTTAACTTGAATCGAG 295
QY	2151	AAGCCCTCTCGATCAAGTCATCGACTCCAGCACTCGGGGATTTTACGACAGGCCAAG 2210
DB	294	ACGTTTTTTTTCGATTAAAGTTTATGATTTTACGATTTACGGGATTTACGATAGAGTTAAG 235
QY	2211	CGAGACCGGATCTGCGCGGCGGACACGCGCGTGTCCGAGATTTGTTCTCCCGCTCAAAGC 2270
DB	234	CGAGAGCGGATTTGTCGTGCTGATACGTTCTGTTCGAGAGTTCGTTTTCTTTTAAAGGC 175
QY	2271	GAGGTGCTGGCCAGCGGGAATCGGAGAAGAACTGCGCALAGGACATTAGCGGCTACAT 2330

Db 174 GAGGTGTTGTTTAAAGCGGATTCGAGAAAGATTGCTTAAGATTAAGCGTTATAT 115
Qy 2331 GAGCTTACTGACAGCTAGACGCGCCCTGCGCCCGCGCCGCGCGCGGAGACC 2330
Db 114 GGTTTTATGCTATAGTATAGTGTCTTTTGTCTTGTGCTTGTGCTGCGGCTGGATT 55
Qy 2331 CGAGCAGCCCTCAGCTCTTCCCGAGCTCGGCTCCGACACTCTCCCTG 2444
Db 54 TTAGTTAGTTTATAGTTTATTTTATTTTATTTTATTTTATTTTATTTT 1

RESULT 13

US-09-918-995-2491
; Sequence 2491, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2491
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-2491

Query Match 13.2%; Score 419; DB 3; Length 457;
Best Local Similarity 98.8%; Pred. No. 3.9e-106;

Matches 422; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2239 AGAATGGGCGCAAGACATTAGCGGCTACTAGCTTCTACTCGACAGCTAGCGCGCC 2358
Db 31 AAACCTGGGCGCAAGACATTAGCGGCTACTAGCTTCTACTCGACAGCTAGCGCGCC 90
Qy 2359 CTGCG 2418
Db 91 CTACCG 150
Qy 2419 GCTCG 2478
Db 151 GCTCG 210
Qy 2479 GCTCG 2538
Db 211 GCTCG 270
Qy 2539 TCTGCAATTAAGCTCAGCGCTTCACTTTGCTGTAACCTTTGTTTCTTACTTACT 2598
Db 271 TCTGCAATTAAGCTCAGCGCTTCACTTTGCTGTAACCTTTGTTTCTTACTTACT 330
Qy 2599 CTTCTTCTGAGATTAATCTCTCTACAAATCCCTCTCCCTCTCTCTCTCTCTCTCT 2658
Db 331 CTTCTTCTGAGATTAATCTCTCTACAAATCCCTCTCCCTCTCTCTCTCTCTCTCT 390
Qy 2659 ACTTCCTCTCTCTGTAATGAATCTCTCACTTTAGAGAGAGCTGGGAGAGTCTGTCAGGC 2718
Db 391 ACTTCCTCTCTCTGTAATGAATCTCTCACTTTAGAGAGAGCTGGGAGAGTCTGTCAGGC 450
Qy 2719 AGCAGCG 2725
Db 451 AGCAGCG 457

RESULT 14
US-10-723-860-2746
; Sequence 2746, Application US/10723860
; Publication No. US20040253606A1

; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUB01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2746
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2746

Query Match 10.9%; Score 346; DB 8; Length 2061;

Best Local Similarity 74.9%; Pred. No. 2.5e-85;

Matches 433; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 912 GCACAGGTGTAACCTGTGCAACAGGCGCTTGTGCTGAATTTACCGGCACCAACGAG 971
Db 799 GCCACGCTAAGCTGTGCAACAGGCGCTTGTGCTGAATTTACCGGCACCAACGAG 858
Qy 972 ATGATCATCAACCAACAGGAGGCGGATGTTCTTTTAACTTAACTTCTGCT 1031
Db 859 ATGATCATCAACCAACAGGAGGCGGATGTTCTTTTAACTTAACTTCTGCT 918
Qy 1032 CTGATCCACAGGCTCAATTAATTTTGTGATGATTTTGTGCGATCCCAATCAC 1091
Db 919 CTCAACCCACCGCCACTACATGTTTGTGGAATGTTCTGCGCACCTTAACAC 978
Qy 1092 TGAAGTTTCAAGGAGCAATGCGTCTTTCGCGCAACCGGACACCAATGTCAGGA 1151
Db 979 TGGCGTTTCAAGGAGCAATGCGTCTTTCGCGCAACCGGACACCAATGTCAGGCG 1038
Qy 1152 AATGGGTCTATATGATCTCGGATTTCCCGCAACCTGGGGCTCATGGAATGCGCAAG 1211
Db 1039 AATGGGTCTATATGATCTCGGATTTCCCGCAACCTGGGGCTCATGGAATGCGCAAG 1098
Qy 1212 ATCTCTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1271
Db 1099 ATCTCTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1158
Qy 1272 ATGATGTTTAACTGCTTTCAGCAATGCAAGCGCGCTGATGTTGTTGAATGTAAC 1331
Db 1159 ATGATGTTTAACTGCTTTCAGCAATGCAAGCGCGCTGATGTTGTTGAATGTAAC 1218
Qy 1332 GAGGAGCGGAGGAGACATGAGCAACCGCGCGCGCTGAGAGAGTTCATCTTCCCTGAG 1391
Db 1219 GAGGAGCGGAGGAGACATGAGCAACCGCGCGCGCTGAGAGAGTTCATCTTCCCTGAG 1278
Qy 1392 ACTGATTCATGCGCGCTGACCGCTTACCAAGAACAGGATTAACAACTGAATAATAGAT 1451
Db 1279 AGCAATTCATGAGAGTACGCTTACCAAAACCGATATTAATCACTTAAGATGAT 1338
Qy 1452 CACAACCTTTTGCAAAAGATTTGCGGATTAATGA 1489
Db 1339 CATTAACCTTTTGCAAAAGGCTTCAGAGACAACTATGA 1376

RESULT 15
US-10-723-860-6839
; Sequence 6839, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &

```

? TITLE OR INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
?
? FILE REFERENCE: 05882.0193.NPUS01
?
? CURRENT APPLICATION NUMBER: US/10/723,860
?
? CURRENT FILING DATE: 2003-11-26
?
? PRIOR APPLICATION NUMBER: 60/429,739
?
? PRIOR FILING DATE: 2002-11-26
?
? NUMBER OF SEQ ID NOS: 8393
?
? SOFTWARE: PatentIn version 3.2
?
? SEQ ID NO 6839
?
? LENGTH: 3137
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
US-10-723-860-6839

```

Query Match	10.8%	Score	342.8	DB	8	Length	3137
Best Local Similarity	74.6%	Pred.	No. 2.5e-84				
Matches	431	Conservative	0	Mismatches	147	Indels	0
						Gaps	0

Oy	912	GCAACAGGCTACCTGTGCACACAGGCCCTTTGGCTGAAATTTCAACCGGCACCAACCGAG	972
Db	817	GCCCAAGCTTACCTGTGCACCGGCCTCTGTGGCTCAAAATTCACCGGACCAAACTGAG	876
Oy	972	ATGATTCATCAACCAACAGGGAGGCGCATGTTCTCTTTTAACTTTTAACTTTCTGCT	1031
Db	877	ATGATTCATTAAGAAACAGGGCGAGGGCATGTTTCTTTCTTGAGCTTCAATATAACGGA	936
Oy	1032	CTGCATCCCAACGGCTCATTAACAATTTTTTGTGATGTGATTTTGGCGGATCCCAATCAC	1091
Db	937	CTCAATCCCACTGGCCACTACAAATGTGTGTGTAAGGTGGTGTGGCGGACCCCAACCCAC	996
Oy	1092	TGGAAGTTCAAGGAGGCAAAATGGGTTCTTTCGGGCAAAAGGAGCAACCAATGTCCAAGA	1151
Db	997	TGGCCCTTCCAGGGGGCAAAATGGGTGACTGTGTGCACAAAGCCGACAAATACATCGAAGGC	1056
Oy	1152	AATCGGCTCTATATGCATCCGGAATCCCCAACAACCTGGGCTCACTGATGCGCCAAAGA	1211
Db	1057	AAACAATAATGTAATGTCAACCCAGAGCTCTTAATCTGGTCCCACTGGATGAGACAGAG	1116
Oy	1212	ATCTCTTTTGGAAATTTAAACTTAACGAAACAACAAAGAGCTTCAAATAACAATGGGAG	1271
Db	1117	ATTTCAATTCGGGAAATTTAAACTCAACAAATAACAAAGGCGCAAAATACAAACAACCCAG	1176
Oy	1272	ATGAGGTTTTTACAGTCCCTGCAACAAGTACAGACCCCGCTGCAATGTGTGGAAGTGAAC	1331
Db	1177	ATGATATGCTTACAAATCTCTTACAAATAACCAACCCCGCATGTGCATATGTGTGAAGTACA	1236
Oy	1332	GAGGACGGCAACGAGACACTAGCCAGCCCGGCCGCTGACAGACGTTCACTTTCCCTGAG	1391
Db	1237	GAGGATGGCGTGGAGAGACTTGAATGAGCCCTCAAAAGACCCAGACTTTTAACTTCTCAGAA	1296
Oy	1392	ACTCAGTTTCATGCCGTCAACGGCTTACCAAGAACACGGAATATTACACACTGAAAAATGAT	1451
Db	1297	ACGCATTTTCATTGCAGTGCATGCTCTTACCAAAAACACCGATATTACTCAACTAAAGTATAT	1356
Oy	1452	CACAACCCCTTTGGCAAAAGGATTTGGGGATTAATTGA	1489
Db	1357	CATAACCCCTTTGGCAAAAGGCTTCAAGAGCACTATGA	1394

Search completed: December 20, 2005, 16:46:37
Job time : 2462 secs

This Page Blank (uspto)